PCT

REQUEST

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

For receiving Office use only

International Application No. PCT/US 01/05287

International Filing Date 28 FEB 2001 (28.02.01)

PCT INTERNATIONAL APPLICATION RO/US

Name of receiving Office and "PCT International Application"

Applicant's or agent's file reference *(if desired) (12 characters maximum)* 1462-PCT-00

Box No. I TITLE OF INVENTION

Expression of an Antimicrobial Peptide Via the Plastid Genome to Control Phytopathogenic Bacteria

Box No. II APPLICANT

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

Auburn Uniiversity
309 Samford Hall,
Auburn University, AL 36849, US

☐ This person is also inventor.

Telephone No.
334-844-4977

Facsimile No.
334-844-5963

Teleprinter No.

State *(that is, country)* of nationality:
US

State *(that is, country)* of residence:
US

This person is applicant for the purposes of: ☐ all designated States ☒ all designated States except the United States of America ☐ the United States of America only ☐ the States indicated in the Supplemental Box

Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

University of Central Florida
4000 Central Florida Blvd.,
Orlando, FL 32816, US

This person is:

☒ applicant only

☐ applicant and inventor

☐ inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:
US

State *(that is, country)* of residence:
US

This person is applicant for the purposes of: ☐ all designated States ☒ all designated States except the United States of America ☐ the United States of America only ☐ the States indicated in the Supplemental Box

☐ Further applicants and/or (further) inventors are indicated on a continuation sheet.

Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE

The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: ☐ agent ☒ common representative

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)*

WEISER, Gerard J.
1600 Market Street, Suite 3600,
Philadelphia, PA 19103-7286, US

Telephone No.
215-751-2427

Facsimile No.
215-568-6946

Teleprinter No.

☐ **Address for correspondence:** Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

Continuation of Box No. III FURTHER APPLICANTS AND/OR (FURTHER) INVENTOR(S)

If none of the following sub-boxes is used, this sheet is not to be included in the request.

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

DANIELL, Henry
1255 Marina Point, #315
Casselberry, FL 32707, US

This person is:

- [ ] applicant only
- [ ] applicant and inventor
- [x] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality: US

State *(that is, country)* of residence: US

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [x] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:

- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [ ] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:

- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [ ] the United States of America only
- [ ] the States indicated in the Supplemental Box

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

This person is:

- [ ] applicant only
- [ ] applicant and inventor
- [ ] inventor only *(If this check-box is marked, do not fill in below.)*

State *(that is, country)* of nationality:

State *(that is, country)* of residence:

This person is applicant for the purposes of:
- [ ] all designated States
- [ ] all designated States except the United States of America
- [ ] the United States of America only
- [ ] the States indicated in the Supplemental Box

- [ ] Further applicants and/or (further) inventors are indicated on another continuation sheet.

Box No.V DESIGNATION OF STATES

The following designations are hereby made under Rule 4.9(a) *(mark the applicable check-boxes; at least one must be marked):*

Regional Patent

☒ **AP ARIPO Patent:** **GH** Ghana, **GM** Gambia, **KE** Kenya, **LS** Lesotho, **MW** Malawi, **MZ** Mozambique, **SD** Sudan, **SL** Sierra Leone, **SZ** Swaziland, **TZ** United Republic of Tanzania, **UG** Uganda, **ZW** Zimbabwe, and any other State which is a Contracting State of the Harare Protocol and of the PCT

☒ **EA Eurasian Patent:** **AM** Armenia, **AZ** Azerbaijan, **BY** Belarus, **KG** Kyrgyzstan, **KZ** Kazakhstan, **MD** Republic of Moldova, **RU** Russian Federation, **TJ** Tajikistan, **TM** Turkmenistan, and any other State which is a Contracting State of the Eurasian

☒ **EP European Patent:** **AT** Austria, **BE** Belgium, **CH and LI** Switzerland and Liechtenstein, **CY** Cyprus, **DE** Germany, **DK** Denmark, **ES** Spain, **FI** Finland, **FR** France, **GB** United Kingdom, **GR** Greece, **IE** Ireland, **IT** Italy, **LU** Luxembourg, **MC** Monaco, **NL** Netherlands, **PT** Portugal, **SE** Sweden, and any other State which is a Contracting State of the European Patent Convention and of the PCT

☒ **OA OAPI Patent:** **BF** Burkina Faso, **BJ** Benin, **CF** Central African Republic, **CG** Congo, **CI** Côte d'Ivoire, **CM** Cameroon, **GA** Gabon, **GN** Guinea, **GW** Guinea-Bissau, **ML** Mali, **MR** Mauritania, **NE** Niger, **SN** Senegal, **TD** Chad, **TG** Togo, and any other State which is a member State of OAPI and a Contracting State of the PCT *(if other kind of protection or treatment desired, specify on dotted line)* .....

National Patent *(if other kind of protection or treatment desired, specify on dotted line):*

☒	**AE**	United Arab Emirates	☒	**LC**	Saint Lucia
☒	**AG**	Antigua and Barbuda	☒	**LK**	Sri Lanka
☒	**AL**	Albania	☒	**LR**	Liberia
☒	**AM**	Armenia	☒	**LS**	Lesotho
☒	**AT**	Austria	☒	**LT**	Lithuania
☒	**AU**	Australia	☒	**LU**	Luxembourg
☒	**AZ**	Azerbaijan	☒	**LV**	Latvia
☒	**BA**	Bosnia and Herzegovina	☒	**MA**	Morocco
☒	**BB**	Barbados	☒	**MD**	Republic of Moldova
☒	**BG**	Bulgaria	☒	**MG**	Madagascar
☒	**BR**	Brazil	☒	**MK**	The former Yugoslav Republic of Macedonia
☒	**BY**	Belarus	☒	**MN**	Mongolia
☒	**BZ**	Belize	☒	**MW**	Malawi
☒	**CA**	Canada	☒	**MX**	Mexico
☒	**CH and LI**	Switzerland and Liechtenstein	☒	**MZ**	Mozambique
☒	**CN**	China	☒	**NO**	Norway
☒	**CR**	Costa Rica	☒	**NZ**	New Zealand
☒	**CU**	Cuba	☒	**PL**	Poland
☒	**CZ**	Czech Republic	☒	**PT**	Portugal
☒	**DE**	Germany	☒	**RO**	Romania
☒	**DK**	Denmark	☒	**RU**	Russian Federation
☒	**DM**	Dominica	☒	**SD**	Sudan
☒	**DZ**	Algeria	☒	**SE**	Sweden
☒	**EE**	Estonia	☒	**SG**	Singapore
☒	**ES**	Spain	☒	**SI**	Slovenia
☒	**FI**	Finland	☒	**SK**	Slovakia
☒	**GB**	United Kingdom	☒	**SL**	Sierra Leone
☒	**GD**	Grenada	☒	**TJ**	Tajikistan
☒	**GE**	Georgia	☒	**TM**	Turkmenistan
☒	**GH**	Ghana	☒	**TR**	Turkey
☒	**GM**	Gambia	☒	**TT**	Trinidad and Tobago
☒	**HR**	Croatia	☒	**TZ**	United Republic of Tanzania
☒	**HU**	Hungary	☒	**UA**	Ukraine
☒	**ID**	Indonesia	☒	**UG**	Uganda
☒	**IL**	Israel	☒	**US**	United States of America
☒	**IN**	India	☒	**UZ**	Uzbekistan
☒	**IS**	Iceland	☒	**VN**	Viet Nam
☒	**JP**	Japan	☒	**YU**	Yugoslavia
☒	**KE**	Kenya	☒	**ZA**	South Africa
☒	**KG**	Kyrgyzstan	☒	**ZW**	Zimbabwe
☒	**KP**	Democratic People's Republic of Korea			
☒	**KR**	Republic of Korea			
☒	**KZ**	Kazakhstan			

Check-boxes reserved for designating States which have become party to the PCT after issuance of this sheet:

☒ CO - Colombia 1

Precautionary Designation Statement: In addition to the designations made above, the applicant also makes under Rule 4.9(b) all other designations which would be permitted under the PCT except any designation(s) indicated in the Supplemental Box as being excluded from the scope of this statement. The applicant declares that those additional designations are subject to confirmation and that any designation which is not confirmed before the expiration of 15 months from the priority date is to be regarded as withdrawn by the applicant at the expiration of that time limit. *(Confirmation (including fees) must reach the receiving Office within the 15-month time limit.)*

Supplemental Box *If the Supplemental Box is not used, this sheet need not be included in the request.*

*1. If, in any of the Boxes, **the space is insufficient** to furnish all the information: in such case, write "Continuation of Box No. ..." [indicate the number of the Box] and furnish the information in the same manner as required according to the captions of the Box in which the space was insufficient, in particular:*

*(i) **if more than two persons are involved as applicants and/or inventors** and no "continuation sheet" is available: in such case, write "Continuation of Box No. III" and indicate for each additional person the same type of information as required in Box No. III. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below;*

*(ii) if, in Box No. II or in any of the sub-boxes of Box No. III, the indication **"the States indicated in the Supplemental Box"** is checked: in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the applicant(s) involved and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is applicant;*

*(iii) if, in Box No. II or in any of the sub-boxes of Box No. III, **the inventor or the inventor/applicant is not inventor for the purposes of all designated States or for the purposes of the United States of America:** in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the inventor(s) and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is inventor;*

*(iv) if, in addition to the agent(s) indicated in Box IV, there are **further agents**: in such case, write "Continuation of Box No. IV" and indicate for each further agent the same type of information as required in Box No. IV;*

*(v) if, in Box No. V, the name of any State (or OAPI) is accompanied by the indication **"patent of addition,"** or **"certificate of addition,"** or if, in Box No. V, the name of the United States of America is accompanied by an indication **"continuation"** or **"continuation-in-part"**: in such case, write "Continuation of Box No. V" and the name of each State involved (or OAPI), and after the name of each such State (or OAPI) , the number of the parent title or parent application and the date of grant of the parent title or filing of the parent application;*

*(vi) if, in Box No. VI, there are **more than three earlier applications whose priority is claimed**: in such case, write "Continuation of Box No. VI" and indicate for each additional earlier application the same type of information as required in Box No. VI;*

*(vii) if, in Box No. VI, **the earlier application is an ARIPO application**: in such case, write "Continuation of Box No. VI", specify the number of the item corresponding to that earlier application and indicate at least one country party to the Paris Convention for the Protection of Industrial Property or one Member of the World Trade Organization for which that earlier application was filed.*

*2. If, with regard to the **precautionary designation statement** contained in Box No. V, the applicant wishes to exclude any State(s) from the scope of that statement: in such case, write "Designation(s) excluded from precautionary designation statement" and indicate the name or two-letter code of each State so excluded.*

*3. If the applicant claims, in respect of any designated Office, the benefits of provisions of the national law concerning **non-prejudicial disclosures or exceptions to lack of novelty**: in such case, write "Statement concerning non-prejudical disclosures or exceptions to lack of novelty" and furnish that statement below.*

Continuation of Box IV:

T. Daniel Christenbury	Reg. No. 31,750	Guy T. Donatiello	Reg. No. 33,167
Paul A. Taufer	Reg. No. 35,703	James A. Drobile	Reg. No. 19,690
Robert A. McKinley	Reg. No. 43,793	Austin R. Miller	Reg. No. 16,602
Sharon Fenick	Reg. No. 45,269	Stewart M. Wiener	Reg. No. 46,20
Joan T. Kluger	Reg. No. 38,940	Michael A. Patane`	Reg. No. 42,982
Felicity Rowe	Reg. No. 47,042	Sharon Fenick	Reg. No. 45,269
Stephenie Yeung	Reg. No. P48.052		

Box No. VI PRIORITY CLAIM ☐ Further priority claims are indicated in the Supplemental Box.

Filing date of earlier application (day/month/year)	Number of earlier application	Where earlier application is: national application: country	regional application:* regional Office	international application: receiving Office
item (1) 29/02/2000 29 FEBRUARY 2000	60/185,662	US		
item (2)				
item (3)				

☒ The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) *(only if the earlier application was filed with the Office which for the purposes of the present international application is the receiving Office)* identified above as item(s): (1)

** Where the earlier application is an ARIPO application, it is mandatory to indicate in the Supplemental Box at least one country party to the Paris Convention for the Protection of Industrial Property for which that earlier application was filed (Rule 4.10(b)(ii)). See Supplemental Box.*

Box No. VII INTERNATIONAL SEARCHING AUTHORITY

Choice of International Searching Authority (ISA) *(if two or more International Searching Authorities are competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used):*

ISA/US

Request to use results of earlier search; reference to that search *(if an earlier search has been carried out by or requested from the International Searching Authority):*

Date *(day/month/year)* Number Country *(or regional Office)*

Box No. VIII CHECK LIST: LANGUAGE OF FILING

This international application contains the following **number of sheets:**

request	:	5
description (excluding sequence listing part)	:	16
claims	:	2
abstract	:	1
drawings	:	8
sequence listing part of description	:	
Total number of sheets	:	32

This international application is **accompanied** by the item(s) marked below:

1. ☒ fee calculation sheet
2. ☐ separate signed power of attorney
3. ☒ copy of general power of attorney; reference number, if any:
4. ☐ statement explaining lack of signature
5. ☐ priority document(s) identified in Box No. VI as item(s):
6. ☐ translation of international application into (language):
7. ☐ separate indications concerning deposited microorganism or other biological material
8. ☐ nucleotide and/or amino acid sequence listing in computer readable form
9. ☒ other *(specify)*: POSTCARD

Figure of the drawings which should accompany the abstract: 2

Language of filing of the international application: English

Box No. IX SIGNATURE OF APPLICANT OR AGENT

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).

Date: 28/2/2001

Gerard J. Weiser

GERARD J. WEISER

(98.02-01)

For receiving Office use only

1. Date of actual receipt of the purported international application: JC09 Rec'd PCT/PTO 28 FEB 2001
2. Drawings: ☐ received: ☐ not received:
3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application:
4. Date of timely receipt of the required corrections under PCT Article 11(2):
5. International Searching Authority (if two or more are competent): ISA/US
6. ☐ Transmittal of search copy delayed until search fee is paid.

For International Bureau use only

Date of receipt of the record copy by the International Bureau:

PCT

GENERAL POWER OF ATTORNEY

(for several international applications filed under the Patent Cooperation Treaty)

(PCT Rule 90.5)

The undersigned person(s):
(Family name followed by given name; for a full legal entity, full official designation. The address must include postal code and name of country.)

Auburn University, 309 Samford Hall, Auburn University, AL 36849, US

hereby appoint(s) the following person as: ☒ agent ☐ common representative

Name and address
(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)

WEISER, Gerard J.	Reg. No. 19,783	Christenbury, T. Daniel	Reg. No. 31,750
Miller, Austin R.	Reg. No. 18,602	Donatiello, Guy T.	Reg. No. 33,187
Drobile, James A.	Reg. No. 16,890	McKinley, Robert A.	Reg. No. 43,793.
Miller, Austin R.	Reg. No. 16,602	Fenick, Sharon	Reg. No. 45,269
Kluger, Joan T.	Reg. No. 38,940	Patane, Michael A.	Reg. No. 42,982
Wiener, Stewart M.	Reg. No. 46,201	Rowe, Felicity	Reg. No. 47,042
Yeung, Stephenie	Reg. No. P48,052		

to represent the undersigned before

☒ all the competent International Authorities

☐ the International Searching Authority only

☐ the International Preliminary Examining Authority only

in connection with any and all international applications filed by the undersigned with the following Office

US as receiving Office

and to make or receive payments on behalf of the undersigned.

Signature(s) *(where there are several persons, each of them must sign; next to each signature, indicate the name of the person signing and the capacity in which the person signs, if such capacity is not obvious from reading this power):*

By: C M Moriarty

Printed Name: C. Michael Moriarty

Title: Associate Provost and Vice President for Research

Date: 2/26/01

Form PCT/Model of general power of attorney (for several international applications) (July 1992) Legalstar 1997, Form PCT-12

PCT

GENERAL POWER OF ATTORNEY

(for several international applications filed under the Patent Cooperation Treaty)

(PCT Rule 90.5)

The undersigned person(s):
(Family name followed by given name; for a full legal entity, full official designation. The address must include postal code and name of country.)

University of Central Florida, 4000 Central Florida Boulevard, Orlando, Florida 32816, US

hereby appoint(s) the following person as: ☒ agent ☐ common representative

Name and address
(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)

WEISER, Gerard J.	Reg . No. 19,763	Christenbury, T. Daniel	Reg. No. 31,750
Miller, Austin R.	Reg. No. 16,602	Donatiello, Guy T.	Reg. No. 33,167
Drobile, James A.	Reg. No. 19,690	McKinley, Robert A.	Reg. No. 43,793.
Miller, Austin R.	Reg. No. 16,602	Fenick, Sharon	Reg. No. 45,269
Kluger, Joan T.	Reg. No. 38,940	Patane, Michael A.	Reg. No. 42,982
Wiener, Stewart M.	Reg. No. 46,201	Rowe, Felicity	Reg. No. 47,042
Yeung, Stephenie	Reg. No. P48,052		

to represent the undersigned before

☒ all the competent International Authorities

☐ the International Searching Authority only

☐ the International Preliminary Examining Authority only

in connection with any and all international applications filed by the undersigned with the following Office

US as receiving Office

and to make or receive payments on behalf of the undersigned.

Signature(s) *(where there are several persons, each of them must sign; next to each signature, indicate the name of the person signing and the capacity in which the person signs, if such capacity is not obvious from reading this power):*

By: _____

Printed Name: Thomas O'Neil

Title: Director of Sponsored Research

Date: 2/26/01

Form PCT/Model of general power of attorney (for several international applications) (July 1992) LegalStar 1997, Form PCTM2

This sheet is not part of and does not count as a sheet of the international application.

PCT

FEE CALCULATION SHEET

Annex to the Request

For receiving Office use only

PCT/US 01/06287

International application No.

Date stamp of the receiving Office

Applicant's or agent's file reference: **1462-PCT-00**

Applicant: **Auburn University and University of Central Florida**

CALCULATION OF PRESCRIBED FEES

1. TRANSMITTAL FEE . . . . . 240.00 T — 240

2. SEARCH FEE . . . . . 700.00 S — 700

International search to be carried out by _____

(If two or more International Searching Authorities are competent in relation to the international application, indicate the name of the Authority which is chosen to carry out the international search.)

3. INTERNATIONAL FEE

Basic Fee

The international application contains 32 sheets.

first 30 sheets . . . . . 382 b1 — 382

2 remaining sheets x $9.00 additional amount = 18.00 b2 — 18

Add amounts entered at b1 and b2 and enter total at B . . 400 ~~18.00~~ B — 400

Designation Fees

The international application contains 87 designations.

6 number of designation fees payable *(maximum 6)* x 82.00 amount of designation fee = 492.00 D — 492

Add amounts entered at B and D and enter total at I . . . . . 892 ~~510.00~~ I — 892

(Applicants from certain States are entitled to a reduction of 75% of the international fee. Where the applicant is (or all applicants are) so entitled, the

4. FEE FOR PRIORITY DOCUMENT *(if applicable)* . . . . . 15.00 P — 15

5. TOTAL FEES PAYABLE . . . . . 1847 ~~1,465.00~~ TOTAL — 1847

Add amounts entered at T, S, I and P, and enter total in the TOTAL box

☐ The designation fees are not paid at this time.

MODE OF PAYMENT

- ☐ authorization to charge deposit account (see below)
- ☐ bank draft
- ☐ coupons
- ☒ cheque
- ☐ cash
- ☐ other *(specify)*:
- ☐ postal money order
- ☐ revenue stamps

DEPOSIT ACCOUNT AUTHORIZATION *(this mode of payment may not be available at all receiving Offices)*

The RO/ US

- ☐ is hereby authorized to charge the total fees indicated above to my deposit account.
- ☒ *(this check-box may be marked only if the conditions for deposit accounts of the receiving Office so permit)* is hereby authorized to charge any deficiency or credit any overpayment in the total fees indicated above to my deposit account.
- ☐ is hereby authorized to charge the fee for preparation and transmittal of the priority document to the International Bureau of WIPO to my deposit account.

13-3405 — Deposit Account No.

28/2/2001 — Date *(day/month/year)*

Signature

09807720

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number
WO 01/64927 A1

(51) International Patent Classification[7]: C12N 15/82, 5/04, 15/10, A01H 5/00

(21) International Application Number: PCT/US01/06287

(22) International Filing Date: 28 February 2001 (28.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/185,662 29 February 2000 (29.02.2000) US

(71) Applicants *(for all designated States except US)*: **AUBURN UNIVERSITY** [US/US]; 309 Samford Hall, Auburn University, AL 36849 (US). **UNIVERSITY OF CENTRAL FLORIDA** [US/US]; 4000 Central Florida Boulevard, Orlando, FL 32816 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **DANIELL, Henry** [US/US]; 1255 Marina Point #315, Casselberry, FL 32707 (US).

(74) Agents: **WEISER, Gerard, J.** et al.; 1600 Market Street, Suite 3600, Philadelphia, PA 19103-7286 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/64927 A1

(54) Title: EXPRESSION OF AN ANTIMICROBIAL PEPTIDE VIA THE PLASTID GENOME TO CONTROL PHYTOPATHOGENIC BACTERIA

(57) Abstract: This invention provides a novel method to confer disease resistance to plants. Plant plastids are transformed using a plastid vector which contains heterologous DNA sequences coding for a cytotoxic antimicrobial peptide. Transgenic plants are capable of fighting off phytopathogenic bacterial infection.

EXPRESSION OF AN ANTIMICROBIAL PEPTIDE VIA THE PLASTID GENOME TO CONTROL PHYTOPATHOGENIC BACTERIA

RELATED APPLICATIONS

This patent application claims the benefit of U.S. Provisional Application No. 60/185,662, filed 2/29/00. This application is here incorporated by reference.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

The work of this invention is supported in part by the USDA-NRICGP grants 95-82770, 97-35504 and 98-0185 to Henry Daniell.

FIELD OF INVENTION

This application pertains to the field of genetic engineering of plant genomes, particularly plastids, and to methods of and engineered plants that express antimicrobial peptides that lead to and result in phytopathogenic bacteria resistance.

DESCRIPTION OF RELATED ART

Zasloff, in U.S. patent 5,643,876 and 4,810,777, entitled "Biologically Active Synthetic Magainin Peptides" and "Antimicrobial Compounds," described a family of synthetic compounds termed "magainin which are capable of inhibiting the growth or proliferation of gram-positive and gram negative bacteria, fungi, virus, and protozoan species.

Haynie, in U. S. patent 5,847,047, entitled "Antimicrobial Composition of Polymer and a Peptide Forming Amphiphilic Helices of the Magainin-Type," offers a series of non-natural oligopeptides that share a common amino acid sequence referred to as the core oligopeptide. Such core oligopeptide has antimicrobial effects. The patent also provides N-addition analogues to the core oligopeptide that exhibit higher antimicrobial effects.

Olsen et. al., in U. S. patent 6,143,498, entitled "Antimicrobial Peptide," proposed a method of producing human antimicrobial peptides from the defensin superfamily through transformation of host cells. Olsen suggested the production of these defensin-related peptides through transformation of host cells with vectors containing the isolated DNA molecules of the peptides.

Kim, et. al., in U. S. patent 6,183,992, entitled "Method For Mass Production Of Antimicrobial Peptide," offered a method of mass producing an antimicrobial peptide. In particular, a fusion gene – containing a basic antimicrovial peptide which ligated directly or

indirectly to a negatively charged acidic peptide having at least two cysteine residues – is cloned into an expression vector targeted toward microorganisms such as E. Coli.

All patents and publications are hereby incorporated by reference in their entireties.

BACKGROUND OF THE INVENTION

Plant diseases caused by bacterial pathogens have had a detrimental effect on global crop production for years. Between 1979 and 1980 India lost up to 60% of its rice crop due to bacterial rice blight. Between 1988 and 1990, there was a 10.1% loss of the global barley crop due to bacterial pathogens, worth $1.9 billion (Baker et al., 1997). In the United States, there was an estimated 44,600 metric ton reduction of soybean crops due to bacterial pathogens in 1994 (Wrath et al., 1996). On the average, pathogens are responsible for a 12-13% reduction of global crop production each year (Dempsey et al., 1998).

A prior effort to combat these devastating pathogens is plant breeding (Mourgues et al., 1998). The results were limited due to the ability of the bacteria to adapt and find a way around the defense mechanism. Agrochemicals have also been used but their application is limited by their toxicity to humans and the environment (Mourgues et al., 1998).

Plant Defense Against Pathogens: Many of the pathways and products in the plant response to phytopathogens have been elucidated with the emergence of molecular biology. The plant defense response can be divided into 3 major categories, early defense (fast), local defense (fast/intermediate) and systemic defense (intermediate to slow) (Mourgues et al., 1998). During the early stage, the plant cell is stimulated by contact with pathogen-produced elicitors. Bacterial genes such as *hrp* (hypersensitive response and pathogenicity) or *avr* (avirulence) genes stimulate the plant defense mechanism (Baker et al., 1997). The most prominent early defense response is the HR (hypersensitive response), which leads to cellular death reducing further infection by the pathogen. Local defense entails cell wall reinforcement, stimulation of secondary metabolite pathways, synthesis of thionins and synthesis of PR (pathogenesis-related) proteins (Mourgues et al., 1998). The final phase is known as SAR (systemic acquired resistance), which protects the uninfected regions of the plant.

Engineering Resistance: Genetic engineering has allowed for some enhancement of natural defense genes from plants by cloning and over-expression in non-host plants. Cloning of resistance (R) genes has been used to protect rice from bacterial leaf blight (Mourgues et al., 1998). Pathogenesis-related (PR) genes have been cloned from barley and have shown to provide resistance to P. syringae pv. *tabaci* (Mourgues et al., 1998). Anti-fungal peptides produced by various organisms have been cloned and studied. However, although anti-fungal development has been promising, bacteria still maintain the ability to adapt to plant defenses.

Those skilled in the art will be familiar with antimicrobial peptides. Examples of some of these substances include PGLa (frog skin), defensins (human phagocytes), cecropins (Silkmoth pupae or pig intestine), apidaecins (honeybee lymph), melittin (bee venom), bombinin (toad skin) and the magainins (frog skin). Specifically bactericidal peptides include large polypeptides such as lysozyme (MW 15000 daltons) and attacins (MW 20-23,000 daltons) as well as smaller polypeptides such as cecropin (MW 4000 daltons) and the magainins (MW 2500 daltons). The spectrum of biocidal activity of these peptides is somewhat correlated to size. In general, the large polypeptides are active against limited types and species of microorganisms (e.g., lysozyme against only gram positive bacteria), whereas many of the smaller oligopeptides demonstrate a broad spectrum of antimicrobial activity, killing many species of both gram positive and gram negative bacteria. It has been shown that magainin, cecropins, and bombinin oligopeptides form similar secondary structures described as an amphiphilic helix (Kaiser et al. Annu. Rev. Biophys. Biophys. Chem 16, 561-581, 1987). These peptides with *a*-helical structures are ubiquitous and found in many organisms. They are believed to participate in the defense against potential microbial pathogens. One of the first biocidal oligopeptides to be isolated from natural sources was bombinin and is described by Csordas et al. (Proc. Int. Symp. Anim. Plant Toxins, 2, 515-523, (1970)). Csordas teaches significant sequence homology between bombinin and melittin, another antimicrobial peptide, isolated from bee venom.

Specifically, the role of magainins from *Xenopus laevis* (African frog) and its analogues have been investigated by Zasloff et al. (WO 9004408) as pharmaceutical compositions such as a broad-spectrum topical agent, a systemic antibiotic; a wound-healing stimulant; and an anticancer agent (Jacob and Zasloff, 1994). Cuervo et al. (WO 9006129) describe the preparation of deletion analogues of magainin I and II for use as pharmaceutical compositions. They disclose a general scheme for the synthetic preparation of compounds with magainin-like activity and structure. However, the possible agricultural use of magainin-type antimicrobial peptides has not yet been explored. Accordingly, it is an objective of this invention to demonstrate the conference of phytopathogenic bacteria resistance to plants by transforming plant cell plastids to express magainin and its analogues.

Plastid Transformation: To date, plastid transformation, particularly has enabled generation of herbicide (Daniell et al., 1998), insect resistant crops (Kota et al., 1999; McBride et al., 1995; DeCosa et al., 2000) and production of pharmaceutical proteins (Guda et al., 2000; Staub et al., 2000). Plastid transformation was selected because of several advantages over nuclear transformation (Daniell, 1999 A, B; Bogorad, 2000; Heifetz, 2000). With concern growing about outcrossing of genetically altered genes, it should be noted that plastid expressed genes are maternally inherited in most crops. Gene containment is possible when foreign genes are engineered via the plastid genome, which prevents pollen transmission in crops that maternally inherit the plastid genome. Because a majority of crop plants inherit their plastid genes maternally, the foreign genes do not escape into the environment.

Although pollen from plants that exhibit maternal inheritance contain metabolically active plastids, the plastid DNA is lost during pollen maturation (Helfetz, 2000). Despite the potential advantage of plastid reproduction of AMPs, it was not obvious that AMPs would be produed in this manner. Prior to the patent application there were no published reports of expression of AMPs in plant plastids.

Non-obviousness of the disease resistance. Several foreign genes have been expressed within plastids to introduce novel traits including herbicide resistance or insect resistance. However, all of these foreign proteins, without exception, function within plastids. For example, herbicides target proteins or enzymes present within plastids. When engineered plastids are consumed by target insects, insecticidal proteins are released inside the insect gut.

However, in order to use the chloroplast compartment to engineer disease resistance, it was necessary to export foreign proteins into the cytosol where phytopathogens colonize. Therefore, it was not obvious to engineer the plastid genome to confer disease resistance. There are no prior reports or suggestions in the literature that plastid genome could be engineered to confer disease resistance. Also, it is known in the art that antimicrobial peptides are toxic to plant chloroplasts because of the charge on the chloroplast membranes. However, this invention teaches that transgenic plastids expressing antimicrobial peptides rupture at the site of infection upon cell death. Release of large amounts of the antimicrobial peptide prevent the spread of the phytopathogen. Thus, the present invention confirms a novel and unobvious solution to combat phytopathogens that is previously unknown and contrary to all current understanding of chloroplast biology.

Most importantly, small peptides are not stable inside living cells and are highly susceptible to proteolytic degradation. For this reason, small peptides are usually produced as fusion proteins with larger peptides in biological systems. Megainin type peptides are chemically synthesized and never made in biological systems for that reason. Therefore, it was not obvious to express a small peptide of a few amino acids within plastids. Successful expression of this antimicrobial peptide was not anticipated but this invention opens the door for expression of several small peptides within plastids, including hormones.

SUMMARY OF THE INVENTION

This invention provides a new option in the battle against phytopathogenic bacteria through transformation of the plant plastid genome. The present invention is applicable to all plastids of plants. These include chromoplasts which are present in the fruits, vegetables and flowers; amyloplasts which are present in tubers like the potato; proplastids in roots; leucoplasts and etioplasts, both of which are present in non-green parts of plants. All known methods of

transformation can be used to introduce the vectors of this invention into target plant plastids including bombardment, PEG Treatment, Agrobacterium, microinjection, etc.

This invention provides plastid expression constructs which are useful for genetic engineering of plant cells and which provide for enhanced expression of a foreign peptide in plant cell plastids. The transformed plant is preferably a metabolically active plastid, such as the plastids found in green plant tissues including leaves and cotyledons. The plastid is preferably one which is maintained at a high copy number in the plant tissue of interest.

The plastid expression constructs for use in this invention generally include a plastid promoter region and a DNA sequence of interest to be expressed in transformed plastids. The DNA sequence may contain one or a number of consecutive encoding regions, one of which preferably encoding an antimicrobial peptide of the magainin family. Plastid expression construct of this invention is linked to a construct having a DNA sequence encoding a selectable marker which can be expressed in a plant plastid. Expression of the selectable marker allows the identification of plant cells comprising a plastid expressing the marker.

In the preferred embodiment, transformation vectors for transfer of the construct into a plant cell include means for inserting the expression and selection constructs into the plastid genome. This preferably comprises regions of homology to the target plastid genome which flank the constructs.

The plastid vector or constructs of the invention preferably include a plastid expression vector which is capable of importing phytopathogenic bacteria resistance to a target plant species which comprises an expression cassette which is described further herein. Such a vector generally includes a plastid promoter region operative in said plant cells' plastids, a DNA sequence which encode at least an antimicrobial peptide of the magainin family. Preferably, expression of one or more DNA sequences of interest will be in the transformed plastids.

The preferred embodiment of the invention provides a universal plastid vector comprising a DNA construct. The DNA construct includes a 5' part of a plastid spacer sequence; a promoter, such as Prrn, which is operative in the plastid of the target plant cells; a heterologous DNA sequence encoding at least one antimicrobial peptide of the magainin family; a gene that confers resistance to a selectable marker such as the aadA gene; a transcription termination region functional in the target plant cells; and flanking each side of the expression cassette, flanking DNA sequences which are homologous to a DNA sequence of the target plastid genome, whereby stable integration of the heterologous coding sequence into the plastid genome of the target plant is facilitated through homologous recombination of the flanking sequence with the homologous sequences in the target plastid genome. The vector may further comprise a ribosome binding site (rbs), a 5' untranslated region (5'UTR). A promoter, such as

psbA, accD or 16srRNA, is to be used in conjunction with the 5'UTR. In addition to the encoding region of the antimicrobial peptide, the heterologous DNA sequence of the DNA construct may also include other genes whose expression are desired.

In another embodiment of the invention, non-universal plastid vectors such as pUC, pBlueScript, pGEM may be used as the agent to insert the DNA construct

This invention provides transformed crops, like solanaceous, monocotyledonous and dicotyledonous plants, that are resistant to phytopathogenic bacteria. Preferably, the plants are edible for mammals, including humans. These plants express an antimicrobial peptide at levels high enough to provide upwards of 96% inhibition of growth against *Pseudomonas syringae*, a major plant pathogen. The transformed plants do not differ morphologically from untransformed plants.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. (A) Chloroplast vector used for transformation of *Nicotiana tabacum* var. Petit Havana. Vector contains the aadA selectable marker gene that confers resistance to spectinomycin, the Prrn promoter, and the TpsbA terminator. (B) Amino acid sequence of the lytic peptide MSI-99.

Figure 2. (A) Phenotype of T_0 and T_1transgenic plants. Plants1-3 are T_0 transgenic plants while plant 4 is untransformed. Plants 5-7 are T_1 transgenic plants. Seedlings germinated on MSO+500μg/ml spectinomycin (B). Three T_1 transgenic lines (1-3) and Control (4).

Figure 3. (A) Primers, 8P and 8M used to confirm integration of foreign genes via PCR. 8P anneals with the 5'end of the aadA gene and 8M anneals with the 3'end of the 16S rDNA gene. PCR analysis of DNA extracted from T_0 (B), T_1 (C) and T_2 (D) plants run on a 0.8% agarose gel. T_0 (B) Lane 1 1kb ladder, 2 through 5 transgenic lines, 6 MSI-99 plasmid. T_1 (C) Lane 1, 1kb ladder, 2 through 4 transgenic, lane 5 plasmid control and lane 6 untransformed plant DNA. T_2 (C) lane 1, 1kb ladder, 2 through 5 transgenic, lane 6 plasmid control and lane 7 untransformed plant DNA.

Figure 4. Southern analysis of T_0 and T_1 generations. (A) Probe used to confirm integration of foreign genes. The 2.3kb probe fragment was cut with BamHI and NotI containing the flanking sequence. (B) Lane 2-6 T_0 transgenic lines, lane 1 untransformed and Lane 7 plasmid DNA. (C) Lanes 2-7 T_1 transgenic lines, Lane 1 untransformed and Lane 8 plasmid DNA.

Figure 5. *In situ* bioassays. 5 to 7mm areas of T_0 transformants and untransformed Petit Havana leaves were scraped with fine grain sandpaper. Ten μl of $8x10^5$, $8x10^4$, $8x10^3$ and $8x10^2$ cells from an overnight culture of P. syringae were added to each prepared area. Photos were taken 5 days after inoculation

Figure 6. *In vitro* bioassays for T_0, T_1 and T_2 generations of 3 transgenic lines (10A, 11A and 13A). Five µl of bacterial cells from an overnight culture were diluted to (A_{600} 0.1-0.3) and incubated for 2 hours at 25°C with 100µg of total plant protein extract. One ml of LB broth was added to each sample. Samples were incubated overnight at temperature appropriate for the specific bacteria. Absorbance at 600nm was recorded. Data was analyzed using GraphPad Prism. Negative control was untransformed plant extract. Buffer only was added as a control and stock culture was used as a reference point.

Figure 7. *In vitro* bioassays for *P. aeruginosa*. Five µl of bacterial cells from an overnight culture were diluted to (A_{600} 0.1-0.3) and incubated for 2 hours at 25°C with 100µg of total protein extract from T_1 plants. One ml of LB broth was added to each sample. Samples were incubated overnight at 37°C. Absorbance at 600nm was recorded. Data was analyzed using GraphPad Prism. Negative control was an untransformed plant extract. Buffer only was added as a control and stock culture was used as a reference point.

Figure 8. Five µl of an overnight culture of *P. syringae* diluted to (A_{600} 0.1-0.3) was mixed with 100µg total protein extract from T2 lines 11A and 13A (germinated in the absence of spectinomycin). After 2-hour incubation, 1ml of LB broth was added to the mixture and incubated over night at 27°C. The following morning absorbance at 600_{nm} was recorded (A). In parallel, 50µl of each mix was plated onto LB plates and incubated overnight at 27°C. The next morning a count of viable CFUs were made using the Bio Rad Gell Dock (B).

DETAILED DESCRIPTION OF THE INVENTION

This invention demonstrates the confering of phytopathogenic resistance in plants through plastid transformation. This invention includes the use of all plastids in plants, including chloroplasts, chloroplasts which are present in fruits, vegetables and flowers, amyloplasts which are present in tubers, proplastids in roots, lencoplasts in non-green parts of plants. In a preferred embodiment of the invention, the chloroplast genome is used. Plastid transformation and expression vectors comprising heterologous DNA encoding magainin and its analogues are provided . The anti-microbial peptide (AMP) used in this invention is an amphipathic alpha-helix molecule that has an affinity for negatively charged phospholipids commonly found in the outer-membrane of bacteria. Upon contact with these membranes, individual peptides aggregate to form pores in the membrane, resulting in bacterial lysis. Because of the concentration dependent action of the AMP, it was expressed via the plastid genome to accomplish high dose delivery at the point of infection. PCR products and Southern blots confirmed plastid integration of the foreign genes and homoplasmy. Growth and development of the transgenic plants was unaffected by expression of the AMP within the plastids. *In vitro* assays with T_0, T_1 and T_2 plants, confirmed the AMP was expressed at levels high enough to provide 86%(T_0), 88%(T_1) and

96%(T_2) inhibition of growth against *Pseudomonas syringae,* a major plant pathogen. *In situ* assays resulted in intense areas of necrosis around the point of infection in control leaves, while transformed leaves showed no signs of necrosis. Even when germinated in the absence of spectinomycin selection, T_2 generation plants showed 96% inhibition of growth against P.syringae.

MSI-99 is an analogue of a naturally occurring peptide (magainin 2) found in the skin of the African frog. Changes have been made to the amino acid sequence to enhance its lytic abilities. Contrary to the prior knowledge in the art which proposed that anti-microbial peptides having high antibacterial activity also have a high potential for toxic activity against the plastid (Everett and Nicholas, 1994), the transgenic plants of this invention grew, flowered and set seeds like the untransformed control.

Key features of cationic peptides such as MSI-99 are a net positive charge, an affinity for negatively charged prokaryotic membrane phospholipids over neutral-charged eukaryotic membranes, and the ability to form aggregates that disrupt the bacterial membrane (Houston et al., 1997; Matsuzaki et al., 1999; Biggin and Sansom, 1999). Given the fact that the outer membrane is an essential and highly conserved part of all bacterial cells, it is highly unlikely that bacteria would be able to adapt (as they have against antibiotics) and to resist the lytic activity of these peptides. In contrast to prokaryotic membranes, the thylakoid membrane consists of primarily glycolipids and galactolipids instead of phospholipids. Monogalactosyldiacylglycerol (MGDG) makes up 50% of membrane lipid and digalactosyldiacylglycerol (DGDG) 30% (Siegenthaler et al., 1998). Both of these lipids are neutral.

An object of this invention is to compartmentalize the expression of the MSI-99 within the plastid. Compartmentalization of lytic enzymes is a natural occurrence in plants. Compartmentalization serves two purposes: to increase the yield of the peptide and to deliver the peptide at the site of the infection. Due to the high copy number associated with plastid expression, a larger amount of the peptide is produced. The higher yield is important due to the concentration-dependent action of the anti-microbial peptide. Further, the peptide would be released at the site of infection during the HR response. When the HR response occurs, cells are lysed. This disrupts the osmotic balance and causes plastids to lyse. This would release the peptide at high concentration resulting in aggregation and formation of pores in the outer membrane of bacteria. This aids in the prevention of the spread of infection by bacteria.

A high level of AMP expression can be expected due to the following reasons. The nature of plastids to move from a somatically unstable heteroplasmic state to a state of homoplasmy itself lends to high expression (Brock and Hagemann, 2000). The A+T % of MSI-99 is 51.39%, which is compatible with the *Nicotiana tobacum* plastid 61% A+T content (Bogorad et al., 1991; Shimada et al., 1991). Also, published reports from our lab report expression of Cry2A operon (A+T content of 65%) at levels as high as 46% total soluble protein (DeCosa et al., 2000).

MSI-99 was most effective against *P. syringae*, evidenced by total inhibition of 1000 *P. syringae* cells with only 1µg/1000 bacteria (Smith et al. unpublished data). Because the lytic activity of antimicrobial peptides is concentration dependent, the amount of antimicrobial peptide required to kill bacteria was used to estimate the level of expression in transgenic plants. Based on the minimum inhibitory concentration, it was estimated that transgenic plants expressed MSI-99 at 21 % of the total soluble protein. Without the availability of antibody for MSI-99, other direct methods of protein estimation were not feasible.

Plastid vectors and plant transformation: The synthetic peptide used in this invention (MSI-99), is an analogue of the naturally occurring 23 amino acid peptide, magainin II. MSI-99 is a 22 amino acid sequence with an overall charge of +6 as shown in Figure 1. The gene cassette used for transformation consisted of the 16S rRNA promoter, the aadA gene, which confers resistance to spectinomycin, the MS1-99 gene and the psbA (photosynthetic binding protein) terminator. The gene construct may contain, in addition to the MSI-99 gene, another heterologous DNA sequence coding for a gene of interest.

Flanking sequences are from the petunia plastid genome as shown in Figure 1A. Transformation efficiency was much lower (7%) than that observed using the pLD vector (91%), which contains tobacco homologous flanking sequences. Other vectors that are capable of plastid transformation may be used to deliver the gene cassette into the plastid genome of the target plant cells. Such vectors do include plastid expression vectors such as pUC, pBlueScript, pGEM, and all others identified by Daniell in US patents number 5,693,507 and 5,932,479. These publications and patents are herein incorporated by reference to the same extent as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. The vectors preferably include a ribosome binding site (rbs) and a 5' untranslated region (5'UTR). A promoter operably in green or non-green plastids is to be used in conjunction with the 5'UTR)

The number of transformants from the total number of shoots determined percent of transformants. Out of 55 spectinomycin resistant shoots screened, only 4 were transformants with the MSI-99 gene and the rest were mutants. All transformants grew healthy with no apparent morphological effects to T_0 and T_1, generations as shown in Figure 2A. T_1, seeds germinated in the presence of spectinomycin produced healthy green seedlings, while control seedlings were bleached as shown in Figure 2B.

Foreign gene integration, homoplasmy and copy number: PCR was performed by landing one primer on the 5'end of the aadA coding sequence, not present in native plastid and the 3'end of the 16S rDNA (Figure 3A). PCR products of T_0, T_1, and T_2 generations yielded the same size product as the plasmid (MSI-99) as shown in Figure 3B,C,D confirming integration of the foreign genes. The probe

used for the Southern analysis was a 2.3kb fragment from the 5'end of the tmI (BamHI) to the 3'end of the 16SrDNA (NotI) (Figure 4A). The plant DNA was digested with BamHI. DNA from untransformed plants produced a 3.269kb fragment and transformed plant DNA produced a 4.65kb fragment. Southern analysis confirmed integration of foreign genes for T_0 and T_1, as shown in Figure 4B,C. Untransformed DNA showed a 3.2kb fragment while the transformed contained a 4.65kb fragment. Presence of some wild type fragments in T_0 transgenic samples indicated some heteroplasmy as shown in Figure 4B. However, DNA from T_1, generation produced only the 4.65Kb fragment confirming homoplasmy. As shown in Figure 4C. A cell is said to be homoplasmic when all of the plastid are uniformly transformed. If only a fraction of the genomes was transformed, the copy number should be less than 10,000 (Bendich, 1987). By confirming that the MSI-99 integrated genome is the only one present in transgenic plants (homoplasmy), one could estimate that the MSI-99 gene copy number could be as many as 10,000 per cell.

Bioassays: T_0 *in situ* assays in potted plants (6 to 7 months old) resulted in areas of necrosis surrounding the point of infection in untransformed control, while transgenic leaves showed no areas of necrosis (Figure 5). Even inoculation of $8X10^5$ cells resulted in no necrosis in transgenic leaves (Figure 5A), suggesting the local concentration of the antimicrobial peptide to be very high. However, untransformed plants inoculated with $8x10^3$ cells displayed intense necrosis as shown in Figure 5B.

Cell free extracts of T_0, T_1, and T_2 transgenic plants displayed a strong ability to inhibit growth of *P. syringae in vitro* by 84%, 86% and 96% compared to untransformed plants as shown in Figure 6. The increase in growth inhibition from T_0 to T_2 can be attributed to heteroplasmy in the T_0 generation that was eliminated in subsequent generations. This indicates the peptides retained their lytic activity and successfully passed on the trait to the subsequent generations. The control had less growth than the buffer only. This is most probably due to natural defense peptides such as defensins and thionins produced by plants (Mourgues et al., 1998). When performing *in vitro* bioassays against *P. aeruginosa*, results were similar with T_1, generation showing 96% inhibition of growth (Figure 7).

Absorbance readings as shown in Figure 8A from transgenic plants germinated in the absence of spectinomycin, displayed 96% inhibition of growth that is comparable to transgenic plants germinated in the presence of spectinomycin. Plated cells of bioassay samples from T2 plants germinated in the absence of spectinomycin as shown in Figure 8B showed 83% inhibition of growth compared to the control. The marginal degree of difference between the plating results and the bioassay results (13%) can be explained by the difference in environment. While the plated bacteria were no longer exposed to active peptides, bacteria in the liquid media were constantly surrounded by active peptides.

Protein Estimation: The plate with 10^{-5} dilution had 43 CFUs. The equated to $43x10^6$ CFU/ml. The count was adjusted to reflect the 5µl of culture used. This resulted in a count of 21,500 bacterial cells in the initial 5µl of culture incubated with the peptide. Using 1 µg to kill 1000 *P. syringae* cells as the reference (Smith et al. unpublished data), the estimated expression of MSI-99 was 21.5µg in 100µg soluble protein (21.5%).

The initial low rate of transformation was most likely due to less than 100% homology between the petunia flanking sequences and the tobacco plastid genome. This is not surprising because very low transformation efficiency was also observed when tobacco plastid flanking sequences were used to transform potato plastid genome (Sidorov et al., 1999). Also, other projects in our lab that use the pLD vector (has tobacco flanking sequences) obtained transformation efficiency of 91% transformants to mutants. T_0 and T_1 transgenic plants were healthy and showed no morphological or developmental abnormalities. Retention of lytic activity was evident in the sharp decrease in bacterial growth in the *in vitro* bioassays (84 to 96%). When comparing Southern blots to lytic activity, lytic activity increased as homoplasmy was reached. Equal lytic activity was also observed in transgenic plants germinated in the absence of spectinomycin (96% inhibition of growth). Transgenic plants transferred to potting soil for 5 to 6 months after being removed from spectinomycin selection, displayed similar antimicrobial properties against inoculations of *P. syringae*. These observations eliminate the possibility that spectinomycin absorbed into the plant tissue during germinationof seeds, may be responsible for the growth inhibition in the *in vitro* and *in situ* bioassays. Also, the observation that MSI-99 was equally active in transgenic plants germinated in the presence or absence of spectinomycin shows the stability of the introduced trait in the absence of any selection pressure.

Plastid expression in crops such as tobacco should allow for mass production of the peptide at a lower cost compared to chemical synthesis or production in *E. coli*. This invention thus demonstrates another option in the on going battle against pathogenic bacteria.

The invention is exemplified by the following non-limiting example.

Example 1

Plant transformation: For plant transformation, Nicotiana tabacum var. Petit Havana seeds were germinated on MSO media at 27^0C with photoperiods of 16 hour light and 8 hour dark. Sterile leaves were bombarded using the Bio-Rad Helium driven PDS-1000/He System. After bombardment, leaves were wrapped and kept in the dark for 48 hours. Leaves were then cut into $1cm^2$ squares and placed on a petri dish containing RMOP media with 500µg/ml spectinomycin (first round of selection). Four to six weeks later, shoots were transferred to fresh media and antibiotic (second round of selection).

Shoots that appeared during the second selection were transferred to bottles containing MSO and spectinomycin (500µg/ml). Plants were screened via PCR for transformation. Those that were PCR positive for the presence of the MSI-99 gene were transferred to pots and grown in chambers at 27°C with photoperiods of 16-hour light and 8-hour dark. After flowering, seeds were harvested and sterilized with a solution of I-part bleach and 2-part water with 1 drop of tween-20. Seeds were vortexed for 5 minutes then washed 6 times with 500µl of dH_20 and dried in speed vac. T^1, and T_2seeds were germinated on MSO + 500µg/ml spectinomycin. Untransformed Petit Havana seeds were germinated on the same media as a control to ensure the spectinomycin was active.

PCR conformation Plant DNA extraction on T_0, T_1, and T_2 was performed using the QIAGEN DNeasy Mini Kit on putative transgenic samples and untransfon-ned plants. PCR primers were designed using Primer Premier software and made by GIBCO BRL. Primer (8p:5'ATCACCGCTTCCCTCATAAATCCCTCCC3') anneals with the 5' end of the aadA and primer (8M:5'CCACCTACAGACGCTTTACGCCCAATCA3') anneals with the 3' end of 16SrDNA as shown in Figure 3. PCR was carried out using the Gene Amp PCR system 2400 (Perkin-Elmer). Samples were run for 29 cycles with the following sequence: 94°C for 1 minute, 65°C for 1 minute and 72°C for 3 minutes. The cycles were proceeded by a 94°C denaturation period and followed by a 72°C final extension period. A 4°C hold followed the cycles. PCR products were separated on agarose gels.

Southern analysis: Integration of foreign genes for T_0 and T_1, was determined by Southern blot analysis. DNA from transformed and untransformed plants was digested with BamHI and run on a 0.7% agarose gel. The DNA was then transferred to a nylon membrane by capillary action. The probe was digested with BamHI and NotI and was labeled with 32 P using the Probe Quant™ G-50 Micro Columnris and protocol (Amersharn). Labeled probe was hybridized with the nylon membrane using the Stratagene QUICK-HYB hybridization solution and protocol. Membrane was exposed to film, and developed.

***In vitro* bioassay**: *P. syringae* and *P.aeruginosa* were cultured overnight prior to the assay. 50 mg of leaf tissue (minus mid-rib) was grounded in a micro-centrifuge containing 150µl of phosphate buffer pH5.5 with 5mM PMSF and 5mM with a plastic pestle. Samples were centrifuged for 5 minutes at 10,000x g at 4°C. Supernatant was transferred to a fresh tube and kept on ice. Protein concentration was determined by Bradford assay. One hundred µg of total plant protein was mixed with 5µl of bacteria from overnight culture in a falcon tube. Initial absorbency ranged from 0.1 to 0.3 (A_{600}). Tubes were incubated for 2 hours at 25°C on a rotary shaker at 125rpm. One ml of LB broth was added and tubes were allowed to incubate for 18 hours at 27°C for *P. syringae* and 37°C for *P. aeruginosa* on a rotary shaker at 125rpm. Absorbance (A_{600}) was read for each tube. Results were statistically analyzed using GraphPad Prism.

To rule out spectinomycin as the cause of growth inhibition, the same experiment with *P. syringae* was repeated using T_2 plants that were geminated on MSO with no spectinomycin. For confirmation of the absorption readings, a serial dilution was made of samples after the initial 2-hour incubation. Dilutions of 10^{-3} to 10^{-5} were plated onto LB plates and incubated overnight at 27°C. The next morning a count of viable CFUs were made using the Bio Rad Gell Dock.

To estimate the level of protein expression, a serial dilution was prepared from the starting bacterial culture ($Absorbance_{600}$, 0.1-0.3) used for the *In vitro* bioassay. Fifty µlof each dilution was plated on LB medium and incubated overnight at 27°C. The following morning, CFUs were counted using the Bio Rad Gel Dock and the amount of cells used in the bioassay was calculated. The minimum inhibitory concentration of Iµg/1000 *P.syringae* cells was used to determine antimicrobial peptide concentration in 100µg of cell free plant extracts.

***In situ* bioassay**: *P. syringae* was cultured overnight prior to the assay. Five to seven mm areas of T_0 transformants and untransformed Petit Havana leaves were scraped with fine grain sandpaper. Ten µl of $8xlO^5$, $8xlO^4$, $8XI0^3$ and $8x10^2$ cells from an overnight culture of P. syringae were added to each prepared area. Photos were taken 5 days after inoculation.

REFERENCES

Baker B, Zambryski P, Staskawicz SP, Dinesh-Kumar. (1997) Signaling in Plant-Microbe Interactions. Science. **276**: 726-723

Bendich A.J. (1987) Why do chloroplasts and mitochondria contain so many copies in their genome? BioEssays. **6**: 279-282.

Biggin P, Sansom M. (1999) Interactions of (x-helices with lipid bilayers: a review of simulation studies. Biophysical Chemistry. **76**: 161-183

Bock R, Hagemann R. (2000) Extracellular inheritance: Plastid genomics: Manipulation of plastid genomes and biotechnological applications. Progress in Botany. **6**: 76-90

Bogorad L. (2000) Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products. TIBTECH **18**: 257-263

Bogarad L, Vasil 1. (1991) The Molecular Biology of Plastids. San Diego: Academic Press, 38

Cary J, Rajasekaran K, Jaynes J, Cleveland T. (2000) Transgenic expression of a gene encoding a synthetic antimicrobial peptide results in inhibition of fungal growth *in vitro* and in planta. Plant Science. **154**: 171-181

Daniell H. (1997) Transformation and foreign expression in plants mediated by microprojectile bombardment. Methods in Molecular Biology. **62**: 463-489

Daniell H. (1999 A) New tools for plastid genetic engineering. Nat. Blotechnol. **17**: 855-856

Daniell H. (1999 B) GM crops: Public perception and scientific solutions. Trends in plant science. **4**: 467-469

Daniell H, Datta R, Varma S, Gray S, Lee S.B. (1998) Containment of herbicide resistance through genetic engineering of the chloroplast genome. Nat. Blotechnol. **16**: 345-348

DeCosa B, Moar W, Lee S.B, Miller M, Daniell H. (2000) HyperExpression of the Cry2Aa2 operon in chloroplasts leads to the formation of insecticidal crystals. Nat. Blotechnol. In press

Everett, Nicholas. (1994) Design of Antifungal peptides for agricultural applications. Ed. Hedin, Paul., Merin, Julius, & Hollingworth, Robert. Washington DC: American Chemical Society, 278-292

Guda C, Lee S.B., Daniell H. (2000) Stable expression of a biodegradable protein-based polymer in tobacco chloroplasts. Plant Cell Reports. **19**: 257-262

Heifetz P. (2000) Genetic engineering of the plastid. Blochimie. **82**: 655-666

Houston Jr M.E., Kondejewski L, Gough M, Fidai S, Hodges R. S, Hancock R. (1997) Influence of performed *a*-helix and *a*-helix induction on the activity of cationic antimicrobial peptides. J.Peptide Res. **52**: 81-88

Jacob L, Zasloff M. (1994) Potential therapeutic applications of magainins and other antimicrobial; agents of animal origin. Antimicrobial peptides. (Ciba Foundation Symposium **186**): 197-223

Kota M, Daniell H, Varma S, Garczynski F, Gould F, Moar WJ. (1999) Over expression of the Bacillus thuringiensis Cry2A protein in plastids confers resistance to plants against susceptible and Btresistant insects. Proc. Natl. Acad.Sci. USA. **96**:1840-1845

Matsuzaki K. (1998) Magainins as paradigm for the mode of action of pore forming polypeptides. Biochimica et Blophysica Acta. **1376**: 391400

McBride K.E, Svab Z, Schaaf D. J, Hogen P.S, Stalker D.M, Maliga P. (I 995) Amplification of a chimeric Bacillus gene in chloroplasts leads to extraordinary level of an insecticidal protein in tobacco. Blo/technology **13**: 362-365

Mourgues F, Brisset M.N, Cheveau E. (1998) Strategies to improve plant resistance to bacterial diseases through genetic engineering. Trends in Biotechnology. **16**: 203-210

Neuhas J, Sticher L, Meins, Jr F, Boller T. (1991) A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. Proc. Natl. Acad.Sci. USA. **88**: 10362-10366

Shimada H, Sugiura M. (1991) Fine structural features of the plastid genome: comparison of the sequenced chloroplast genomes. Nucleic Acids Research **19**: 983-995

Sidorov V, Kasten D, Pang S, Hajdukiewicz P, Saub J, Nehra N. (1999) Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. The Plant Journal. **19**: 209-216

Siegenthaler Paul-Andr6, Murata Noria. Ed. (1998) Lipids in Photosynthesis: Structure, Function and Genetics. Boston: Kluwer, 1-52, 119-144.

Staub J, Garcia B, Graves J, Hajdukiewicz P, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll J, Saptola L, Ward D, Ye G, Russell D. (2000) High-yield production of a human therapeutic protein in tobacco chloroplasts. Nat. Biotechnol. **18**: 333-338.

Tümnieler B, Kiewitz C. (1999) Cystic Fibrosis: an inherited ibility to bacterial infections. Molecular Medicine Today. **5**: 3 5 1–358

Utsugi T, Schroit A, Connor J, Bucana C, Fidler L. (1991) Elevated expression of phosphatidylserine in the outer leaflet of human tumor cells and recognition by activated human blood monocytes. Cancer Res. **51**: 3062-3066

Wrath J.A, Anderson T. R, Arsyad D.M, Gai J, Ploper L. D, Portapuglia A, Ram H. H, Yorinori J T.. (1996) Soybean Disease Loss Estimates for the Top 10 Soybean Producing Counties in 1994. Plant Disease. **81**: 107-110

What is claimed is:

1. A stable plastid transformation and expression vector which comprises an expression cassette comprising, as operably linked components in the 5' to the 3' direction of translation, a promoter operative in said plastid, a selectable marker sequence, a heterologous DNA sequence coding for cytotoxic antimicrobial peptide (AMP), transcription termination functional in said plastid, and flanking each side of the expression cassette, flanking DNA sequences which are homologous to a DNA sequence of the target plastid genome, whereby stable integration of the heterologous coding sequence into the plastid genome of the target plant is facilitated through homologous recombination of the flanking sequence with the homologous sequences in the target plastid genome.

2. A vector of claim 1, wherein the plastid is selected from the group consisting of chloroplasts, chromoplasts, amyloplasts, proplastide, leucoplasts and etioplasts.

3. A vector of claim 1, wherein the antimicrobial peptide is selected from the groups of defensins, PGLA (frog skin), cecropins, apidaecins, melittin, bombinin and magainin.

4. A vector of claim 3, wherein the antimicrobial peptide is magainin I or II.

5. A vector of claim 1, wherein the selectable marker sequence is an antibiotic-free selectable marker.

6. A universal integration and expression vector of claim 1 competent for stably transforming a plastid genome of different plant species wherein the flanking DNA sequences are homologous to a spacer sequence of the target plastid genome and the sequence is conserved in the plastid genome of different plant species.

7. A stably transformed plant which comprises plastid stably transformed with the vector of claims 1, 2, 3, 4, 5 or 6 or the progeny thereof, including seeds.

8. A stably transformed plant of claim 7 which is a solanaceous plant.

9. A stably transformed plant of claim 7 which is a monocotyledonous or dicotyledonous plant.

10. A stably transformed plant of claim 9 which is maize, rice, grass, rye, barley, oat, wheat, soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

11. A stably transformed plant of claim 7 which is edible for mammals and humans.

12. A stably transformed plant of claim 7 in which all the chloroplasts are uniformly transformed.

13. A stably transformed plant of claim 7 in which the transformed plastid of the plants including subsequent generations are capable of enhanced levels of expression.

14. A stably transformed plant of claim 7 in which transgenic plants germinated in the absence of antibiotic selectable marker sequence, like spectinomycin.

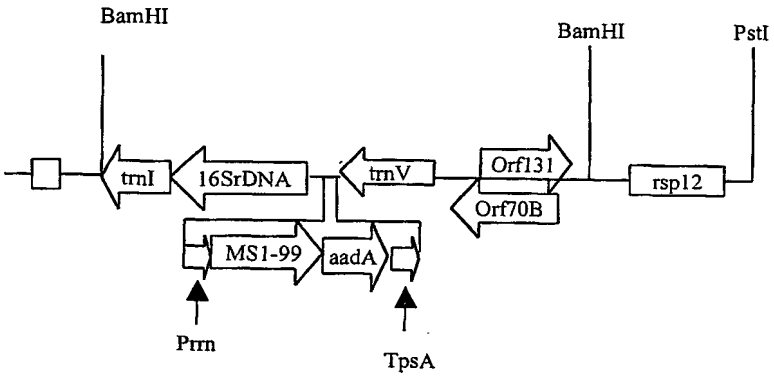
15. A method for stably transforming a target plant to control a phytopathogenic bacteria which comprises introducting an integration and expression vector of claims 1, 2, 3, 4, 5 or 6 into a plastid genome of the target plant, and allowing the transformed plant to grow.

16. A vector of any one of claims 1 – 14, wherein the antimicrobial peptide is a cationic amphiphipathic alpha-helix molecule which has affinity for negatively charged phospholipides in the outer membrane of the target bacteria and which is functional to form aggregates that disrupt and lyse the bacterial membrane of the target microbe, and in the prevention of the spread of infection by the bacteria.

17. A vector of any one of claims 1-14, wherein said vector further comprises a ribosome binding site (rbs) and a 5' untranslated region (5'UTR).

18. A method of claim 15, wherein said vector further comprises a ribosome binding site (rbs) and a 5' untranslated region (5'UTR).

A

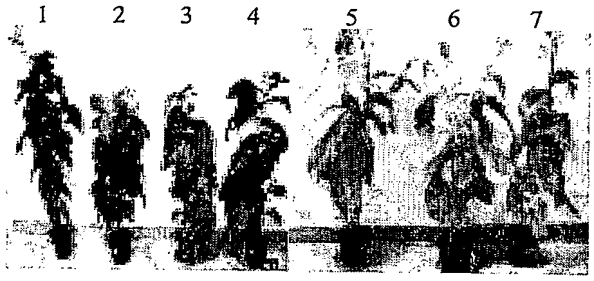


B

MS1-99:
GIGKFLKSAKKFGKAFVKILNS
(+6 charge)

Figure 1

A



T_0 T_1

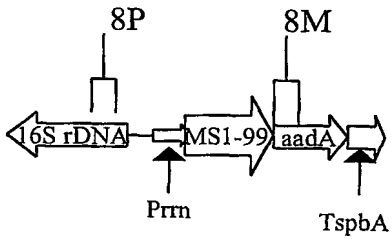
B

Figure 2

A

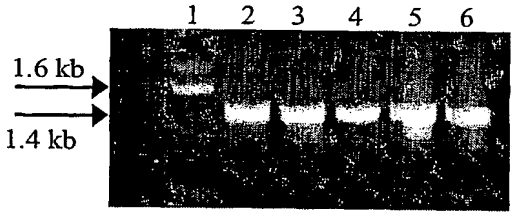
8P : 5' CCACCTACAGACGCTTTACGCCCAATCA 3'

8M: 5' ATCACCGCTTCCCTCATAAATCCCTCCC 3'



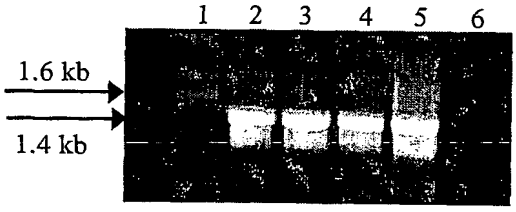
B

T0



C

T1



D

T2

Figure 3

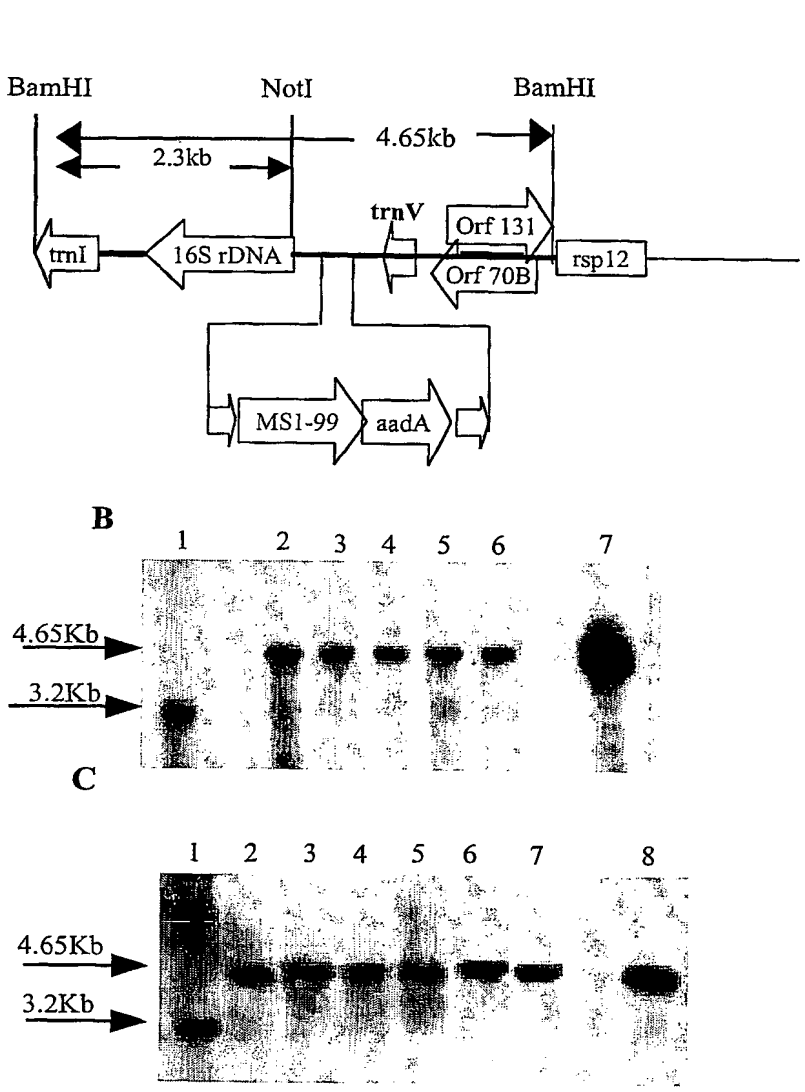

A
BamHI
NotI
BamHI
4.65kb
2.3kb
trnV
Orf 131
trnI
16S rDNA
Orf 70B
rsp12
MS1-99
aadA
B
1 2 3 4 5 6 7
4.65Kb
3.2Kb
C
1 2 3 4 5 6 7 8
4.65Kb
3.2Kb

Figure 4

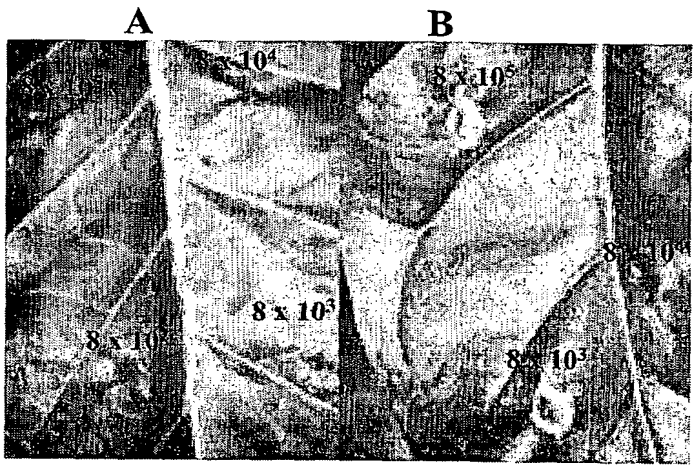
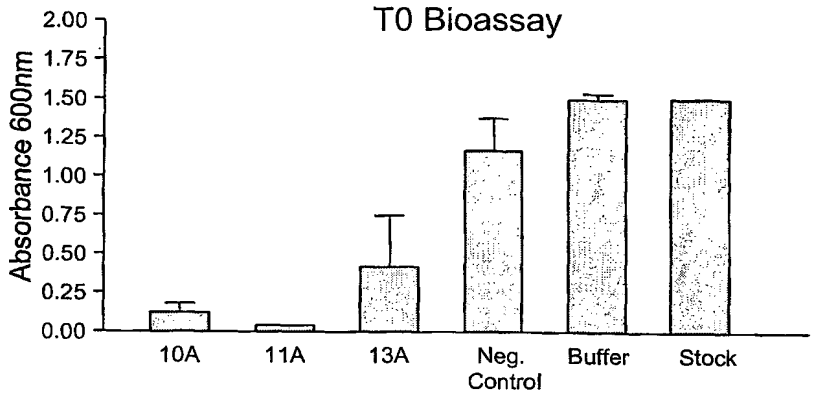
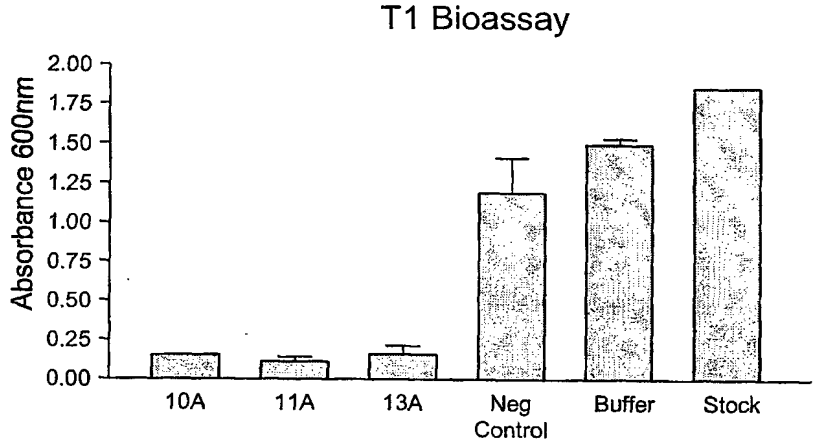


Figure 5

A


T0 Bioassay
Absorbance 600nm
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
10A
11A
13A
Neg. Control
Buffer
Stock

B


T1 Bioassay
Absorbance 600nm
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
10A
11A
13A
Neg Control
Buffer
Stock

C

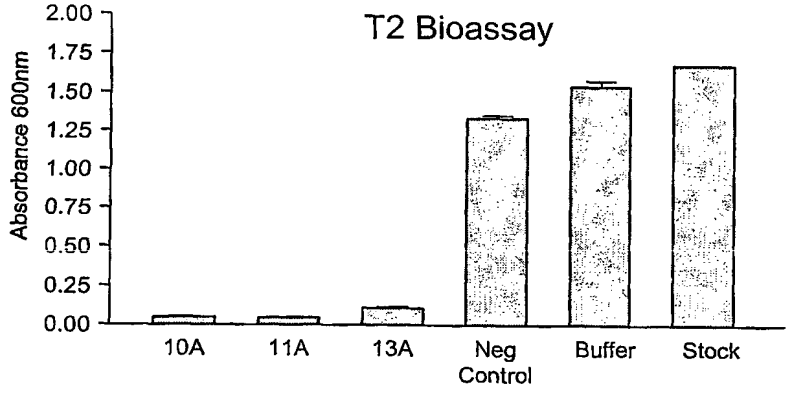

T2 Bioassay
Absorbance 600nm
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
10A
11A
13A
Neg Control
Buffer
Stock

Figure 6

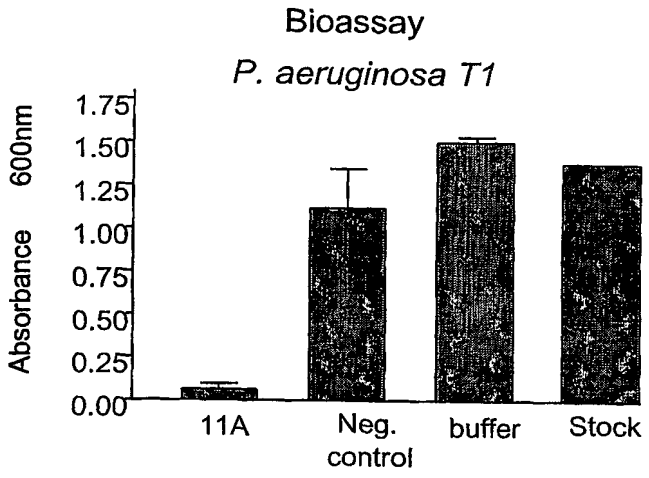
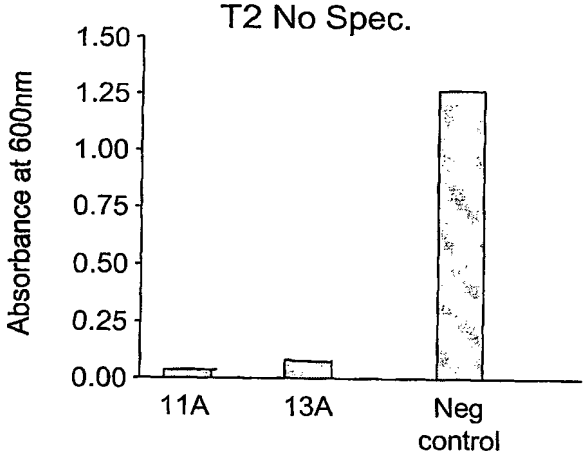

Bioassay
P. aeruginosa T1
Absorbance 600nm
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
11A
Neg. control
buffer
Stock

Figure 7

A


T2 No Spec.
Absorbance at 600nm
1.50
1.25
1.00
0.75
0.50
0.25
0.00
11A
13A
Neg control

B

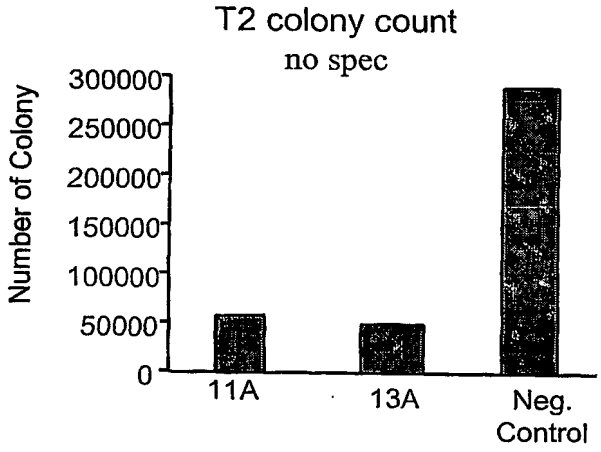

T2 colony count
no spec
Number of Colony
300000
250000
200000
150000
100000
50000
0
11A
13A
Neg. Control

Figure 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/06287

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :C12N 15/82, 5/04, 15/10; A01H 5/00
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,877,402 A (MALIGA et al) 02 March 1999, col. 17, lines 9-58, col. 25, lines 4-67, col. 27, lines 34-41, col. 29, line 27, to col. 30, line 31, Fig. 5, Fig. 6C-E, Fig. 18B, col. 7, lines 17-36, col. 7, lines 54-67, col. 11, lines 9-23.	1-15, 18
Y	US 5,451,513 A (MALIGA et al) 19 September 1995, col. 9, lines 35-48, col. 10, lines 25-38.	1-15, 18
Y	WO 99/06564 A1 (SANFORD SCIENTIFIC, INC.) 11 February 1999, page 2, lines 3-28, page 5, lines 3-11, page 9, lines 23-27, page 18, lines 4-29.	1-15, 18

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 20 APRIL 2001	Date of mailing of the international search report 23 MAY 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer ANNE R. KUBELIK JOYCE A. BRIDGERS PARALEGAL SPECIALIST TECHNOLOGY CENTER 1600 Telephone No. (703) 308-0198

Form PCT/ISA/210 (second sheet) (July 1998)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/06287

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 90/11770 A1 (CALGENE, INC.) 18 October 1990, page 4, lines 21-28, page 11, line 32, to page 15, line 14, page 19, lines 8-18.	1-15, 18

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/06287

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 16-17
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/06287

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

800/279, 288, 301, 317, 320.1, 320.2, 320, 320.3, 312, 317.3, 317.4, 314, 317.2; 435/320.1, 468, 418, 419, 411, 412, 414, 415, 417

B. FIELDS SEARCHED
Minimum documentation searched
Classification System: U.S.

800/279, 288, 301, 317, 320.1, 320.2, 320, 320.3, 312, 317.3, 317.4, 314, 317.2; 435/320.1, 468, 418, 419, 411, 412, 414, 415, 417

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

AGRICOLA, BIOSIS CABA, CAPLUS, USPAT, EPO, JPO, DERWENT
search terms: (plastid or chloroplast) transformation, defensin, PGLA, cecropin, apidaecin, melettin, bombinin, magainin, antimicrobial

09/807720
1621

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 1462-PCT-00	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/US01/06287	International filing date *(day/month/year)* 28 FEBRUARY 2001	Priority date *(day/month/year)* 29 FEBRUARY 2000
International Patent Classification (IPC) or national classification and IPC Please See Supplemental Sheet.		
Applicant AUBURN UNIVERSITY		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 8 sheets.

   [X] This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 2 sheets.

3. This report contains indications relating to the following items:

   - I [X] Basis of the report
   - II [ ] Priority
   - III [ ] Non-establishment of report with regard to novelty, inventive step or industrial applicability
   - IV [X] Lack of unity of invention
   - V [X] Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
   - VI [ ] Certain documents cited
   - VII [X] Certain defects in the international application
   - VIII [X] Certain observations on the international application

Date of submission of the demand 21 SEPTEMBER 2001	Date of completion of this report 19 AUGUST 2002
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Joyce Bridges ANNE R. KUBELIK Telephone No. (703) 308-0196

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US01/06287

I. Basis of the report

1. With regard to the **elements** of the international application:*

[ ] the international application as originally filed

[X] the description:
pages 1-16
pages NONE , as originally filed
pages NONE , filed with the demand
, filed with the letter of

[X] the claims:
pages NONE
pages NONE , as originally filed
pages 17-18 , as amended (together with any statement) under Article 19
pages NONE , filed with the demand
, filed with the letter of

[X] the drawings:
pages 1-8
pages NONE , as originally filed
pages NONE , filed with the demand
, filed with the letter of

[X] the sequence listing part of the description:
pages NONE
pages NONE , as originally filed
pages NONE , filed with the demand
, filed with the letter of

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language _____ which is:

[ ] the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).

[ ] the language of publication of the international application (under Rule 48.3(b)).

[ ] the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

[ ] contained in the international application in printed form.

[ ] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readable form.

[ ] The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] The statement that the information recorded in computer readable form is identical to the writen sequence listing has been furnished.

4. [X] The amendments have resulted in the cancellation of:

[X] the description, pages NONE

[X] the claims, Nos. NONE

[X] the drawings, sheets/~~fig~~ NONE

5. [ ] This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

* *Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).*

** *Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.*

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US01/06287

IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

[X] restricted the claims.

[ ] paid additional fees.

[ ] paid additional fees under protest.

[ ] neither restricted nor paid additional fees.

2. [ ] This Authority found that the requirement of unity of invention is not complied with and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

[ ] complied with.

[X] not complied with for the following reasons:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s)1-7 and 9-20, drawn to a method of producing plants resistant to plant pathogenic bacteria by transformation of their plastids with vectors encoding antimicrobial peptides.

Group II, claim(s) 8, drawn to the antimicrobial peptide MSI-99.

The inventions listed as Groups I-II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Groups I and II do not share a special technical feature.

The technical feature shared by the groups is a sequence encoding an antimicrobial protein or the antimicrobial protein itself. Smith et al (WO 99/06564) teach the antimicrobial peptide MSI-99 (pg 10-12), rendering claim 8 not novel.

Thus, the technical feature shared by the Groups does not constitute an advance over the prior art.

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

[ ] all parts.

[X] the parts relating to claims Nos. 1-7 and 9-20.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US01/06287

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. statement

Novelty (N)	Claims	1-7 and 9-20	YES
	Claims	NONE	NO
Inventive Step (IS)	Claims	NONE	YES
	Claims	1-7 and 9-20	NO
Industrial Applicability (IA)	Claims	1-7 and 9-20	YES
	Claims	NONE	NO

2. citations and explanations (Rule 70.7)

Claims 1-7 and 9-20 meet the criteria set out in PCT Article 33(4), because methods of producing plants resistant to pathogenic bacteria have industrial applicability.

Claims 1-7 and 9-20 meet the criteria set out in PCT Article 33(2), because the prior art does not teach a method of producing plants resistant to pathogenic bacteria by transformation of their plastids with vectors encoding antimicrobial peptides.

Claims 1-5, 7 and 9-20 lack an inventive step under PCT Article 33(3) as being obvious over MALIGA et al (1999, US Patent 5,877,402) in view of SANFORD SCIENTIFIC, INC. (WO 99/06564).

The claims are drawn to plastid transformation vectors encoding antimicrobial proteins and a method of using those vectors to transform plastids to produce plants that control a plant pathogenogenic bacteria.

MALIGA et al teach disclose methods of tobacco plastid transformation and vectors comprising a promoter, a selectable marker sequence, a gene of interest, a transcription termination sequence, and flanking DNA sequences that are homologous to DNA sequences in the target plastid genome; these vectors also have ribosome binding sites and 5' UTRs (column 7, lines 17-36 and 54-67; column 11, lines 9-23; column 17, lines 9-58; column 25, lines 4-67; column 29, line 27, to column 30, line 31; Fig. 5; Fig. 6C-E; Fig. 18E). Maliga et al do not disclose an antimicrobial peptide gene as that gene of interest.

SANFORD SCIENTIFIC, INC. teach plants transformed with genes encoding the antimicrobial peptides magainin 2, PGL, cecropin B and analogues thereof (pg 12-18).

At the time the invention was made, it would have been obvious to one of ordinary skill in the art to modify the method of transforming plant plastids as taught by MALIGA et al, to express the antimicrobial peptides described in SANFORD SCIENTIFIC, INC. One of ordinary skill in the art would have been motivated to do so because of the suggestion of SANFORD (Continued on Supplemental Sheet.)

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

Claims 1-18 are objected to under PCT Rule 66.2(a)(iii) as containing the following defect(s) in the form or contents thereof:

Claims 2-6, 8-14 and 16-18 start with improper articles.
In claim 15, "introducting" in line 2 is misspelled.
In claim 15, there is an improper article before "integration" in line 2.
An article is missing before "cytotoxic" in claim 1, line 4; "transcription" in claim 1, line 4; "plastid" in claim 7, line 1; "transgenic" in claim 14, line 1; and "antibiotic" in claim 14, line 2.

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

Claims 1-3 and 5-18 are objected to under PCT Rule 66.2(a)(v) as lacking clarity under PCT Article 6 because the claims are indefinite for the following reason(s):

Claim 1 lacks antecedent basis for the limitations "said plastid" in line 3, "the target plastid genome" in line 6, "the heterologous coding sequence" in line 7, "the plastid genome" in line 7, "the target plant" in line 7, "the flanking sequence" in line 8 and "the homologous sequences" in lines 8-9.

It is unclear in claim 1 what is means for a vector to be stable and how a stable vector differs from an unstable one.

Claim 1 is indefinite in its recitation of "transcription termination functional" in lines 4-5. It is unclear what a transcription termination is and in what manner it would be functional. A word, possibly --sequence--, appears to be missing after "termination".

In claim 2, because "plastid" in line 1 is singular, to be grammatically consistent all members of the groups must be singular.

Claim 3 lacks antecedent basis for the limitation "the antimicrobial peptide" in line 1

Claim 3 is not written in proper Markush format. The claim should be in the format "selected from the group consisting of A, B, C and D." in line 2, "groups" should be replaced with – group consisting--. Additionally, as "peptide" in line 1 is singular, all members of the groups should also be singular to be grammatically consistent.

Claim 3 is indefinite in its recitation of the abbreviation "PGLA", which is not defined in the description.

In claim 4, "antimicrobial peptide" should be replaced with --magainin-- or the claim should be made dependent upon claim 1.

Claim 5 is indefinite in its recitation of "wherein the selectable marker sequence is an antibiotic-free selectable marker". Applicant appears to be equating a DNA sequence with what it encodes.

Claim 6 lacks antecedent basis for the limitation "A universal integration and expression vector of claim 1".

Claim 6 "vector ... competent for stably transforming a plastid genome". The phrase --wherein the vector is-- should be inserted before "competent". It is also unclear what it means for a vector to be competent for stably transforming something.

It is unclear in claim 6 what the phrase starting with "wherein" is intended to modify. It currently modifies "species". It is suggested that --and-- be inserted before "wherein".

(Continued on Supplemental Sheet.)

Supplemental Box
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

CLASSIFICATION:

The International Patent Classification (IPC) and/or the National classification are as listed below:
IPC(7): C12N 15/12, 15/82, 15/90; A01H 5/00 and US Cl.: 800/279, 288, 301, 317, 320.1, 320.2, 320, 320.3, 312, 317.3, 317.4, 314, 317.2, 306; 435/320.1, 468, 418, 419, 411, 412, 414, 415, 417

V. 2. REASONED STATEMENTS - CITATIONS AND EXPLANATIONS (Continued):
SCIENTIFIC, INC. to express these genes in plastids (pg 9, lines 9-13) and the suggestion of MALIGA et al to express in plastids genes that confer disease resistance (column 27, lines 34-41).

Claims 1-7 and 9-20 lack an inventive step under PCT Article 33(3) as being obvious over the prior art as applied in the immediately preceding paragraph and further in view of CALGENE, INC. (WO 90/11770).

The claims are drawn to plastid transformation vectors encoding antimicrobial proteins, including magainin I, magainin II, and defensins, and a method of using those vectors to transform plastids to produce plants that control a plant pathogenic bacteria.

MALIGA et al in view of SANFORD SCIENTIFIC, INC. disclose a method of producing plants resistant to pathogenic bacteria by transformation of their plastids with vectors encoding magainin 2, PGL, cecropin B and analogues thereof. MALIGA et al in view of SANFORD SCIENTIFIC, INC. do not disclose the use of other antimicrobial peptide genes in those vectors.

CALGENE, INC. teach plants transformed with genes encoding other antimicrobial peptides, including magainin I (pg 24-50).

At the time the invention was made, it would have been obvious to one of ordinary skill in the art to modify the method of producing pathogen-resistant plants by transformation of their plastids with genes encoding antimicrobial peptides as taught by MALIGA et al in view of SANFORD SCIENTIFIC, INC., to use other antimicrobial peptide genes as described in CALGENE, INC. One of ordinary skill in the art would have been motivated to do so because use of different antimicrobial peptide genes is an obvious design choice.

In a response filed 27 February 2002, Applicant urges that the art at the time of the instant invention believed that high antimicrobial activity would be toxic for the plastid, that any antimicrobial peptide made in the plastid would be degraded by proteolysis, and that no sequences for exporting proteins from the chloroplast were known. Applicant urges that the instant invention overcomes these problems by exporting the foreign proteins from the plastids and that plastid proteins are released upon the damage caused by infection. Applicant urges that none of the cited references teaches flanking sequences conserved among higher plants.

This is not found persuasive because the claims are not drawn to vectors that can export proteins from the plastids, and the instant description does not teach such vectors. The suggestion of Sanford Scientific et al to express antimicrobial peptide genes in the plastids argues against the art believing that antimicrobial peptides would be toxic to plastids or degraded by them. Any protein expressed in a plastid would be released when cells are damaged by infection.

---------- NEW CITATIONS ----------
NONE

VIII. CERTAIN OBSERVATIONS ON THE APPLICATION (Continued):

It is unclear in claim 6 how many different plant species the vector is "competent" for transforming and the sequence is conserved within.

Claim 6 lacks antecedent basis for the limitations "the flanking DNA sequences" and "the sequence". Claims 1 and 6 refer to many different sequences, and it is not clear to which "the sequence" refers.

Claim 7 is indefinite in its recitation of "the progeny thereof". It is not clear if the progeny are of the vectors of claims 1-6, the plastid or the plant.

In claims 7 and 13, the word "including" renders the claims indefinite because it is unclear whether the limitations following the word are part of the claimed invention.

In claims 8-11, "which" should be replaced with --, wherein the plant--.

In claim 10, an article should be inserted after "is" and --plant-- should be inserted after "cotton".

In claim 11, it is unclear what is meant that a plant is "edible for mammals and humans" as any plant can be eaten.

Supplemental Box
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

Additionally, it is not clear why humans are mentioned separately, as they are mammals.

In claims 12-14, "in which" should be replaced with --, wherein --

Claim 12 lacks antecedent basis for the limitation "all the chloroplasts".

It is unclear in claim 12, what it means for chloroplasts to be uniformly transformed.

Claim 13 lacks antecedent basis for the limitation "the transformed plastid of the plants". Additionally, it appears that all the plants share a single plastid.

In claim 13, it is unclear which gene has enhanced levels of expression or if it even is a gene that has enhanced levels of expression. The claim is also indefinite in its recitation of "capable of" as it is not clear if enhanced levels of expression are required.

Claim 14 appears to be missing words and makes no sense.

In claim 14, the word "like" renders the claim indefinite because it is unclear whether the limitations following the word are part of the claimed invention. Additionally, spectinomycin is not a marker sequence.

Claims 15 and 18 are indefinite because they lack agreement between the preamble of the methods and the positive method steps. Methods must be circular; the final step must generate the item the method is intended to produce. For example, the method of stably transforming a plant to control phytopathogenic bacteria in claim 15 ends in allowing a plant to grow, when it should end in the control of phytopathogenic bacteria.

In claim 15, "which" should be replaced with --, wherein the method--. As it stands, the phrase starting with "which" modifies "bacteria".

Claim 15 lacks antecedent basis for the limitations "an integration and expression vector of claims 1, 2, 3, 4, 5 or 6" in lines 2-3 and "the transformed plant" in line 3.

Claims 16-17 are indefinite for being improper multiple dependent claims.

Claim 16 lacks antecedent basis for the limitations "the target bacteria" in line 3 and "the target microbe" in line 4 and "the bacteria" in line 5.

The "functional to form aggregates" and the "functional ... in the prevention" language in lines 3-5 of claim 16 is awkward and ungrammatical.

It is not clear in claims 17 and 18 where the rbs and the 5' UTR are located relative to the other components of the vector.

What is claimed is:

1. A stable plastid transformation and expression vector which comprises an expression cassette comprising, as operably linked components in the 5' to the 3' direction of translation, a promoter operative in said plastid, a selectable marker sequence, a heterologous DNA sequence coding for cytotoxic antimicrobial peptide (AMP), transcription termination functional in said plastid, and flanking each side of the expression cassette, flanking DNA sequences which are homologous to a DNA sequence of the target plastid genome, whereby stable integration of the heterologous coding sequence into the plastid genome of the target plant is facilitated through homologous recombination of the flanking sequence with the homologous sequences in the target plastid genome.

2. A vector of claim 1, wherein the vector further comprises a ribosome binding site (RBS) and a 5' untranslated region (5'UTR).

3. A vector of claim 1, wherein the plastid is selected from the group consisting of chloroplasts, chromoplasts, amyloplasts, proplastide, leucoplasts and etioplasts.

4. A vector of claim 1, wherein the antimicrobial peptide is a cationic amphiphipathic alpha-helix molecule which has affinity for negatively charged phospholipides located in an outer membrane of a target microbe and which is functional to form aggregates that disrupt and lyse the microbe membrane of the target microbe, and in preventing spreading of infection by the microbe.

5. A vector of claim 1, wherein the antimicrobial peptide is selected from the groups of defensins, PGLA (frog skin), cecropins, apidaecins, melittin, bombinin and magainins.

6. A vector of claim 1, wherein the antimicrobial peptide is magainin I or an analogue thereof.

7. A vector of claim 1, wherein the antimicrobial peptide is magainin II or an analogue thereof.

8. The antimicrobial peptide of claim 7, wherein the analogue of magainin II is MSI-99.

9. A vector of claim 1, wherein the selectable marker sequence is an antibiotic-free selectable marker.

10. A vector of claim 1 which is a universal vector competent for stably transforming a plastid genome of different plant species wherein flanking DNA sequences are homologous to a spacer sequence of a target plastid genome and the sequences are conserved in the plastid genome of different plant species.

11. A stably transformed plant which comprises plastid stably transformed with the vector of claim 1, and the progeny or subsequent generations thereof, including seeds.

12. A stably transformed plant of claim 11 which is a solanaceous plant.

13. A stably transformed plant of claim 11 which is a monocotyledonous or dicotyledonous plant.

14. A stably transformed plant of claim 13 which is maize, rice, grass, rye, barley, oat, wheat, soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

15. A stably transformed plant of claim 11 which is edible for mammals and humans.

16. A stably transformed plant of claim 11 in which all the plastids are uniformly transformed.

17. The stably transformed plant and progeny of claim 11, wherein subsequent generations are capable of enhanced levels of expression of the cytotoxic antimicrobial peptide.

18. A stably transformed plant of claim 11 in which transgenic plants germinated in the absence of antibiotic selectable marker sequence, like spectinomycin.

19. A method for stably transforming a target plant to control a phytopathogenic bacteria which comprises transforming plastids of said target plant with a transformation and expression vector of claim 1, selecting transformed plant cells, and allowing the transformed plant cells to grow.

20. A method of claim 19, wherein said vector further comprises a ribosome binding site (RBS) and a 5' untranslated region (5'UTR).

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number
WO 01/64927 A1

(51) International Patent Classification[7]: C12N 15/82, 5/04, 15/10, A01H 5/00

(21) International Application Number: PCT/US01/06287

(22) International Filing Date: 28 February 2001 (28.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/185,662 29 February 2000 (29.02.2000) US

(71) Applicants *(for all designated States except US)*: **AUBURN UNIVERSITY** [US/US]; 309 Samford Hall, Auburn University, AL 36849 (US). **UNIVERSITY OF CENTRAL FLORIDA** [US/US]; 4000 Central Florida Boulevard, Orlando, FL 32816 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **DANIELL, Henry** [US/US]; 1255 Marina Point #315, Casselberry, FL 32707 (US).

(74) Agents: **DONATIELLO, Guy, T.** et al.; Schnader Harrison Segal & Lewis, LLP, 1600 Market Street - Suite 3600, Philadelphia, PA 19103 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with amended claims*

Date of publication of the amended claims: 7 February 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/64927 A1

(54) Title: EXPRESSION OF AN ANTIMICROBIAL PEPTIDE VIA THE PLASTID GENOME TO CONTROL PHYTOPATHOGENIC BACTERIA

(57) Abstract: This invention provides a novel method to confer disease resistance to plants. Plant plastids are transformed using a plastid vector which contains heterologous DNA sequences coding for a cytotoxic antimicrobial peptide. Transgenic plants are capable of fighting off phytopathogenic bacterial infection.

AMENDED CLAIMS

[received by the International Bureau on 23 July 2001 (23.07.01);
original claims 1-18 replaced by amended claims 1-20 (2 pages)]

1. A stable plastid transformation and expression vector which comprises an expression cassette comprising, as operably linked components in the 5' to the 3' direction of translation, a promoter operative in said plastid, a selectable marker sequence, a heterologous DNA sequence coding for cytotoxic antimicrobial peptide (AMP), transcription termination functional in said plastid, and flanking each side of the expression cassette, flanking DNA sequences which are homologous to a DNA sequence of the target plastid genome, whereby stable integration of the heterologous coding sequence into the plastid genome of the target plant is facilitated through homologous recombination of the flanking sequence with the homologous sequences in the target plastid genome.

2. A vector of claim 1, wherein the vector further comprises a ribosome binding site (RBS) and a 5' untranslated region (5'UTR).

3. A vector of claim 1, wherein the plastid is selected from the group consisting of chloroplasts, chromoplasts, amyloplasts, proplastide, leucoplasts and etioplasts.

4. A vector of claim 1, wherein the antimicrobial peptide is a cationic amphiphipathic alpha-helix molecule which has affinity for negatively charged phospholipides located in an outer membrane of a target microbe and which is functional to form aggregates that disrupt and lyse the microbe membrane of the target microbe, and in preventing spreading of infection by the microbe.

5. A vector of claim 1, wherein the antimicrobial peptide is selected from the groups of defensins, PGLA (frog skin), cecropins, apidaecins, melittin, bombinin and magainins.

6. A vector of claim 1, wherein the antimicrobial peptide is magainin I or an analogue thereof.

7. A vector of claim 1, wherein the antimicrobial peptide is magainin II or an analogue thereof.

8. The antimicrobial peptide of claim 7, wherein the analogue of magainin II is MSI-99.

9. A vector of claim 1, wherein the selectable marker sequence is an antibiotic-free selectable marker.

10. A vector of claim 1 which is a universal vector competent for stably transforming a plastid genome of different plant species wherein flanking DNA sequences are homologous to a spacer sequence of a target plastid genome and the sequences are conserved in the plastid genome of different plant species.

11. A stably transformed plant which comprises plastid stably transformed with the vector of claim 1, and the progeny or subsequent generations thereof, including seeds.

12. A stably transformed plant of claim 11 which is a solanaceous plant.

13. A stably transformed plant of claim 11 which is a monocotyledonous or dicotyledonous plant.

14. A stably transformed plant of claim 13 which is maize, rice, grass, rye, barley, oat, wheat, soybean, peanut, grape, potato, sweet potato, pea, canola, tobacco, tomato or cotton.

15. A stably transformed plant of claim 11 which is edible for mammals and humans.

16. A stably transformed plant of claim 11 in which all the plastids are uniformly transformed.

17. The stably transformed plant and progeny of claim 11, wherein subsequent generations are capable of enhanced levels of expression of the cytotoxic antimicrobial peptide.

18. A stably transformed plant of claim 11 in which transgenic plants germinated in the absence of antibiotic selectable marker sequence, like spectinomycin.

19. A method for stably transforming a target plant to control a phytopathogenic bacteria which comprises transforming plastids of said target plant with a transformation and expression vector of claim 1, selecting transformed plant cells, and allowing the transformed plant cells to grow.

20. A method of claim 19, wherein said vector further comprises a ribosome binding site (RBS) and a 5' untranslated region (5'UTR).

09/807720

Copy for the Elected Office (EO/US)

PCT/US01/06287

PATENT COOPERATION TREATY

PCT

From the INTERNATIONAL BUREAU

To:

DONATIELLO, Guy, T.
Schnader Harrison Segal & Lewis, LLP
1600 Market Street - Suite 3600
Philadelphia, PA 19103
ETATS-UNIS D'AMERIQUE

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

Date of mailing (day/month/year)
20 novembre 2001 (20.11.01)

Applicant's or agent's file reference
1462-PCT-00

IMPORTANT NOTIFICATION

International application No.
PCT/US01/06287

International filing date (day/month/year)
28 février 2001 (28.02.01)

1. The following indications appeared on record concerning:

[ ] the applicant [ ] the inventor [X] the agent [ ] the common representative

Name and Address
WEISER, Gerard, J.
1600 Market Street, Suite 3600
Philadelphia, PA 19103-7286
United States of America

State of Nationality

State of Residence

Telephone No.
215-751-2427

Facsimile No.
215-568-6946

Teleprinter No.

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

[X] the person [ ] the name [ ] the address [ ] the nationality [ ] the residence

Name and Address
DONATIELLO, Guy, T.
Schnader Harrison Segal & Lewis, LLP
1600 Market Street - Suite 3600
Philadelphia, PA 19103
United States of America

State of Nationality

State of Residence

Telephone No.
(215) 751-2463

Facsimile No.
(215) 972-7238

Teleprinter No.

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

[X] the receiving Office
[ ] the International Searching Authority
[X] the International Preliminary Examining Authority
[ ] the designated Offices concerned
[X] the elected Offices concerned
[ ] other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer
Sean Taylor

Telephone No.: (41-22) 338.83.38

Form PCT/IB/306 (March 1994)

004468568

09/806,750

PCT/US01/06287

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Commissioner
US Department of Commerce
United States Patent and Trademark Office, PCT
2011 South Clark Place Room CP2/5C24
Arlington, VA 22202
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 20 November 2001 (20.11.01)	
International application No. PCT/US01/06287	**Applicant's or agent's file reference** 1462-PCT-00
International filing date (day/month/year) 28 February 2001 (28.02.01)	**Priority date** (day/month/year) 29 February 2000 (29.02.00)
Applicant DANIELL, Henry	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

21 September 2001 (21.09.01)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer Sean Taylor
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

Form PCT/IB/331 (July 1992)

US0106287

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/82, 15/10, C12P 21/02

A1

(11) International Publication Number: **WO 98/31823**

(43) International Publication Date: 23 July 1998 (23.07.98)

(21) International Application Number: PCT/US98/00840

(22) International Filing Date: 16 January 1998 (16.01.98)

(30) Priority Data:

60/035,955	17 January 1997 (17.01.97)	US
60/069,400	12 December 1997 (12.12.97)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/035,955 (CIP)
Filed on	17 January 1997 (17.01.97)
US	60/069,400 (CIP)
Filed on	12 December 1997 (12.12.97)

(71) Applicant *(for all designated States except US)*: THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: MAYFIELD, Stephen [US/US]; 1238 Sea Village Drive, Cardiff, CA 92007 (US).

(74) Agents: HOLMES, Emily et al.; The Scripps Research Institute, 10550 North Torrey Pines Road, TPC–8, La Jolla, CA 92037 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR EXPRESSION OF RECOMBINANT MOLECULES

(57) Abstract

The present invention relates to a gene expression system in eukaryotic and prokaryotic cells, preferably plant cells and intact plants. In particular, the invention relates to an expression system having a RB47 binding site upstream of a translation initiation site for regulation of translation mediated by binding of RB47 protein, a member of the poly(A) binding protein family. Regulation is further effected by RB60, a protein disulfide isomerase. The expression system is capable of functioning in the nucleus/cytoplasm of cells and in the chloroplast of plants. Translational regulation of a desired molecule is enhanced approximately 100 fold over that obtained without RB47 binding site activation.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR EXPRESSION OF RECOMBINANT MOLECULES

Technical Field

The invention relates to expression systems and methods for expression of desired genes and gene products in cells. Particularly, the invention relates to a gene encoding a RNA binding protein useful for regulating gene expression in cells, the protein binding site, a gene encoding a regulating protein disulfide isomerase and methods and systems for gene expression of recombinant molecules.

Background

Expression systems for expression of exogenous foreign genes in eukaryotic and prokaryotic cells are basic components of recombinant DNA technology. Despite the abundance of expression systems and their wide-spread use, they all have characteristic disadvantages. For example, while expression in *E. coli* is probably the most popular as it is easy to grow and is well understood, eukaryotic proteins expressed therein are not properly modified. Moreover, those proteins tend to precipitate into insoluble aggregates and are difficult to obtain in large amounts. Mammalian expression systems, while practical on small-scale protein production, are more difficult, time-consuming and expensive than in *E. coli*.

A number of plant expression systems exist as well as summarized in US Patent No. 5,234,834, the disclosures of which are hereby incorporated by reference. One advantage of plants or algae in an expression system is that they can be used to produce pharmacologically important proteins and enzymes on a large scale and in relatively pure form. In addition, micro-algae have several unique characteristics that make them ideal organisms for the production of proteins on a large scale.

First, unlike most systems presently used to produce transgenic proteins, algae can be grown in minimal media (inorganic salts) using sunlight as the energy source. These algae can be grown in contained fermentation vessels or on large scale in monitored ponds. Ponds of up to several acres are routinely used for the production of micro-algae. Second, plants and algae have two distinct compartments, the cytoplasm and the chloroplast, in which proteins can be expressed. The cytoplasm of algae is similar to that of other eukaryotic organisms used for protein expression, like yeast and insect cell cultures. The chloroplast is unique to plants and algae and proteins expressed in this environment are likely to have properties different from those of cytoplasmically expressed proteins.

The present invention describes an expression system in which exogenous molecules are readily expressed in either prokaryotic or eukaryotic hosts and in either the cytoplasm or chloroplast. These beneficial attributes are based on the discovery and cloning of components of translation regulation in plants as described in the present invention.

Protein translation plays a key role in the regulation of gene expression across the spectrum of organisms (Kozak, Ann. Rev. Cell Biol., 8:197-225 (1992) and de Smit and Van Duin, Prog. Nucleic Acid Res. Mol. Biol., 38:1-35 (1990)). The majority of regulatory schemes characterized to date involve translational repression often involving proteins binding to mRNA to limit ribosome association (Winter et al., Proc. Natl. Acad. Sci., USA, 84:7822-7826 (1987) and Tang and Draper, Biochem., 29:4434-4439 (1990)). Translational activation has also been observed (Wulczyn and Kahmann, Cell, 65:259-269 (1991)), but few of the underlying molecular mechanisms for this type of regulation have been identified. In plants, light

activates the expression of many genes. Light has been shown to activate expression of specific chloroplast encoded mRNAs by increasing translation initiation (Mayfield et al., Ann. Rev. Plant Physiol. Plant Mol. Biol., 46:147-166 (1995) and Yohn et al., Mol. Cell Biol., 16:3560-3566 (1996)). Genetic evidence in higher plants and algae has shown that nuclear encoded factors are required for translational activation of specific chloroplast encoded mRNAs (Rochaix et al., Embo J., 8:1013-1021 (1989), Kuchka et al., Cell, 58:869-876 (1989), Girard-Bascou et al., Embo J., 13:3170-3181 (1994), Kim et al., Plant Mol. Biol., 127:1537-1545 (1994).

In the green algae *Chlamydomonas reinhardtii*, a number of nuclear mutants have been identified that affect translation of single specific mRNAs in the chloroplast, often acting at translation initiation (Yohn et al., supra, (1996)). Mutational analysis of chloroplast mRNAs has identified sequence elements within the 5' untranslated region (UTR) of mRNAs that are required for translational activation (Mayfield et al., supra, (1995), Mayfield et al., J. Cell Biol., 127:1537-1545 (1994) and Rochaix, Ann. Rev. Cell Biol., 8:1-28 (1992)), and the 5' UTR of a chloroplast mRNA can confer a specific translation phenotype on a reporter gene *in vivo* (Zerges and Rochaix, Mol. Cell Biol., 14:5268-5277 (1994) and Staub and Maliga, Embo J., 12:601-606 (1993).

Putative translational activator proteins were identified by purifying a complex of four proteins that binds with high affinity and specificity to the 5' UTR of the chloroplast encoded *psb*A mRNA [encoding the D1 protein, a major component of Photosystem II (PS II)] (Danon and Mayfield, Embo J., 10:3993-4001 (1991)). Binding of these proteins to the 5' UTR of *psb*A mRNA correlates with translation of this mRNA under a variety of

physiological (Danon and Mayfield, id., (1991)) and biochemical conditions (Danon and Mayfield, Science, 266:1717-1719 (1994) and Danon and Mayfield, Embo J., 13:2227-2235 (1994)), and in different genetic backgrounds (Yohn et al., supra, (1996)). The binding of this complex to the *psb*A mRNA can be regulated *in vitro* in response to both redox potential (Danon and Mayfield, Science, 266:1717-1719 (1994)) and phosphorylation (Danon and Mayfield, Embo J., 13:2227-2235 (1994)), both of which are thought to transduce the light signal to activate translation of *psb*A mRNA. The 47 kDa member of the *psb*A RNA binding complex (RB47) is in close contact with the RNA, and antisera specific to this protein inhibits binding to the *psb*A mRNA *in vitro* (Danon and Mayfield, supra, (1991)).

Although the translational control of psbA mRNA by RB47 has been reported, the protein has not been extensively characterized and the gene encoding RB47 has not been identified, cloned and sequenced. In addition, the regulatory control of the activation of RNA binding activity to the binding site by nuclear-encoded *trans*-acting factors, such as RB60, have not been fully understood. The present invention now describes the cloning and sequencing of both RB47 and RB60. Based on the translation regulation mechanisms of RB47 and RB60 with the RB47 binding site, the present invention also describes a translation regulated expression system for use in both prokaryotes and eukaryotes.

Brief Description of the Invention

The RB47 gene encoding the RB47 activator protein has now been cloned and sequenced, and the target binding site for RB47 on messenger RNA (mRNA) has now been identified. In addition, a regulatory protein disulfide isomerase, a 60 kilodalton protein

referred to as RB60, has also been cloned, sequenced and characterized. Thus, the present invention is directed to gene expression systems in eukaryotic and prokaryotic cells based on translational regulation by RB47 protein, its binding site and the RB60 regulation of RB47 binding site activation.

More particularly, the present invention describes the use of the RB47 binding site, i.e., a 5' untranslated region (UTR) of the chloroplast *psb*A gene, in the context of an expression system for regulating the expression of genes encoding a desired recombinant molecule. Protein translation is effected by the combination of the RB47 binding site and the RB47 binding protein in the presence of protein translation components. Regulation can be further imposed with the use of the RB60 regulatory protein disulfide isomerase. Therefore, the present invention describes reagents and expression cassettes for controlling gene expression by affecting translation of a coding nucleic acid sequence in a cell expression system.

Thus, in one embodiment, the invention contemplates a RB47 binding site sequence, i.e., a mRNA sequence, typically a mRNA leader sequence, which contains the RB47 binding site. A preferred RB47 binding site is *psb*A mRNA. For use in expressing recombinant molecules, the RB47 binding site is typically inserted 5' to the coding region of the preselected molecule to be expressed. In a preferred embodiment, the RB47 binding site is inserted into the 5' untranslated region along with an upstream *psb*A promoter to drive the expression of a preselected nucleic acid encoding a desired molecule. In alternative embodiments, the RB47 binding site is inserted into the regulatory region downstream of any suitable promoter present in a eukaryotic or prokaryotic expression vector. Preferably, the RB47 binding site is positioned within 100 nucleotides of the

translation initiation site. In a further aspect, 3' to the coding region is a 3' untranslated region (3' UTR) necessary for transcription termination and RNA processing.

Thus, in a preferred embodiment, the invention contemplates an expression cassette or vector that contains a transcription unit constructed for expression of a preselected nucleic acid or gene such that upon transcription, the resulting mRNA contains the RB47 binding site for regulation of the translation of the preselected gene transcript through the binding of the activating RB47 protein. The RB47 protein is provided endogenously in a recipient cell and/or is a recombinant protein expressed in that cell.

Thus, the invention also contemplates a nucleic acid molecule containing the sequence of the RB47 gene. The nucleic acid molecule is preferably in an expression vector capable of expressing the gene in a cell for use in interacting with a RB47 binding site. The invention therefore contemplates an expressed recombinant RB47 protein. In one embodiment, the RB47 binding site and RB47 encoding nucleotide sequences are provided on the same genetic element. In alternative embodiments, the RB47 binding site and RB47 encoding nucleotide sequences are provided separately.

The invention further contemplates a nucleic acid molecule containing the sequence encoding the 69 kilodalton precursor to RB47. In alternative embodiments, the RB47 nucleic acid sequence contains a sequence of nucleotides to encode a histidine tag. Thus, the invention relates to the use of recombinant RB47, precursor RB47, and histidine-modified RB47 for use in enhancing translation of a desired nucleic acid.

The invention further contemplates a nucleic acid molecule containing a nucleotide sequence of a polypeptide which

regulates the binding of RB47 to RB47 binding site. A preferred regulatory molecule is the protein disulfide isomerase RB60. The RB60-encoding nucleic acid molecule is preferably in an expression vector capable of expressing the gene in a cell for use in regulating the interaction of RB47 with a RB47 binding site. Thus, the invention also contemplates an expressed recombinant RB60 protein. In one embodiment, the RB47 binding site, RB47 encoding and RB60 encoding nucleotide sequences are provided on the same genetic element. In alternative embodiments, the expression control nucleotide sequences are provided separately. In a further aspect, the RB60 gene and RB47 binding site sequence are provided on the same construct.

The invention can therefore be a cell culture system, an *in vitro* expression system or a whole tissue, preferably a plant, in which the transcription unit is present that contains the RB47 binding site and further includes a (1) transcription unit capable of expressing RB47 protein or (2) the endogenous RB47 protein itself for the purpose of enhancing translation of the preselected gene having an RB47 binding site in the mRNA. Preferred cell culture systems are eukaryotic and prokaryotic cells. Particularly preferred cell culture systems include plants and more preferably algae.

A further preferred embodiment includes (1) a separate transcription unit capable of expressing a regulatory molecule, preferably RB60 protein, or (2) the endogenous RB60 protein itself for the purpose of regulating translation of the preselected gene having an RB47 binding site in the mRNA. In an alternative preferred embodiment, one transcription unit is capable of expressing both the RB47 and RB60 proteins. In a further aspect, the RB47 binding site sequence and RB60 sequence are provided on the same construct.

In one aspect of the present invention, plant cells endogenously containing RB47 and RB60 proteins are used for the expression of recombinant molecules, such as proteins or polypeptides, through activation of the RB47 binding in an exogenously supplied expression cassette. Alternatively, stable plant cell lines containing endogenous RB47 and RB60 are first generated in which RB47 and/or RB60 proteins are overexpressed. Overexpression is obtained preferably through the stable transformation of the plant cell with one or more expression cassettes for encoding recombinant RB47 and RB60. In a further embodiment, stable cell lines, such as mammalian or bacterial cell lines, lacking endogenous RB47 and/or RB60 proteins are created that express exogenous RB47 and/or RB60.

Plants for use with the present invention can be a transgenic plant, or a plant in which the genetic elements of the invention have been introduced. Based on the property of controlled translation provided by the combined use of the RB47 protein and the RB47 binding site, translation can be regulated for any gene product, and the system can be introduced into any plant species. Similarly, the invention is useful for any prokaryotic or eukaryotic cell system.

Methods for the preparation of expression vectors is well known in the recombinant DNA arts, and for expression in plants is well known in the transgenic plant arts. These particulars are not essential to the practice of the invention, and therefore will not be considered as limiting.

The invention allows for high level of protein synthesis in plant chloroplasts and in the cytoplasm of both prokaryotic and eukaryotic cells. Because the chloroplast is such a productive plant organ, synthesis in chloroplasts is a preferred site of translation by virtue of the large amounts of protein that can

be produced. This aspect provides for great advantages in agricultural production of mass quantities of a preselected protein product.

The invention further provides for the ability to screen for agonists or antagonists of the binding of RB47 to the RB47 binding site using the expression systems as described herein. Antagonists of the binding are useful in the prevention of plant propagation.

Also contemplated by the present invention is a screening assay for agonists or antagonists of RB60 in a manner analogous to that described above for RB47. Such agonists or antagonists would be useful in general to modify expression of RB60 as a way to regulate cellular processes in a redox manner.

Kits containing expression cassettes and expression systems, along with packaging materials comprising a label with instructions for use, as described in the claimed embodiments are also contemplated for use in practicing the methods of this invention.

Other uses will be apparent to one skilled in the art in light of the present disclosures.

Brief Description of Drawings

In the figures forming a portion of this disclosure:

Figures 1A-1D show the complete protein amino acid residue sequence of RB47 is shown from residues 1-623, together with the corresponding nucleic acid sequence encoding the RB47 sequence, from base 1 to base 2732. The nucleotide coding region is shown from base 197-2065, the precursor form. The mature form is from nucleotide position 197-1402. Also shown is the mRNA leader, bases 1-196, and poly A tail of the mRNA, bases 2066-2732. Both the nucleotide and amino acid sequence are listed in SEQ ID NO

5.

Figures 2A-2B show the complete protein amino acid residue sequence of RB60 is shown from residues 1-488, together with the corresponding nucleic acid sequence from base 1 to base 2413, of which bases 16-1614 encode the RB60 sequence. Both the nucleotide and amino acid sequence are listed in SEQ ID NO 10.

Figures 3A-3C show the complete sequence of the psbA mRNA, showing both encoded psbA protein amino acid residue sequence (residues 1-352) and the nucleic acid sequence as further described in Example 3 is illustrated. Both the nucleotide and amino acid sequence are listed in SEQ ID NO 13.

Figure 4 is a schematic diagram of an expression cassette containing on one transcription unit from 5' to 3', a promoter region derived from the *psb*A gene for encoding the D1 protein from *C. reinhardtii* further containing a transcription initiation site (TS), the RB47 binding site, a region for insertion of a foreign or heterologous coding region, a RB47 coding region, a RB60 coding region, and the 3' flanking region containing transcription termination site (TS), flanked by an origin of replication and selection marker. Restriction endonuclease sites for facilitating insertion of the independent genetic elements are indicated and further described in Example 4A.

Figures 5A-5B show the nucleotide and amino acid sequence of the RB47 molecule containing a histidine tag, the sequences of which are also listed in SEQ ID NO 14.

Figure 6 is a schematic diagram of an expression cassette containing on one transcription unit from 5' to 3', a promoter region derived from the *psb*A gene for encoding the D1 protein from *C. reinhardtii* further containing a transcription initiation site (TS), the RB47 binding site, a region for

insertion of a foreign or heterologous coding region, a RB47 coding region, and the 3' flanking region containing transcription termination site (TS). Restriction endonuclease sites for facilitating insertion of the independent genetic elements are indicated and further described in Example 4E.

Figure 7 is a schematic diagram of an expression cassette containing on one transcription unit from 5' to 3', a promoter region derived from the *psb*A gene for encoding the D1 protein from *C. reinhardtii* further containing a transcription initiation site (TS), the RB47 binding site, a region for insertion of a foreign or heterologous coding region, and the 3' flanking region containing transcription termination site (TS). Restriction endonuclease sites for facilitating insertion of the independent genetic elements are indicated and further described in Example 4G.

Figure 8 is a Western blot of a tetanus toxin single chain antibody expressed with a construct of the present invention as further described in Example 4G1).

Figure 9 is a schematic diagram of an expression cassette containing on one transcription unit from 5' to 3', a promoter region derived from the *psb*A gene for encoding the D1 protein from *C. reinhardtii* further containing a transcription initiation site (TS), the RB47 binding site, a region for insertion of a coding sequence of bacterial luciferase A and B proteins including the translation termination codon TAA. The 3' flanking region containing transcription termination site (TS). Restriction endonuclease sites for facilitating insertion of the independent genetic elements are indicated and further described in Example 4G2).

Figure 10 illustrates the accumulation of expressed bacterial luciferase protein in the chloroplast as further

described in Example 4G2).

Figure 11 is a schematic diagram of an expression cassette containing on one transcription unit from 5' to 3', a promoter region derived from the *psb*A gene for encoding the D1 protein from *C. reinhardtii* further containing a transcription initiation site (TS), the RB47 binding site, a region for insertion of a foreign or heterologous coding region for dimeric IgA (dIgA) and the 3' flanking region containing transcription termination site (TS). Restriction endonuclease sites for facilitating insertion of the independent genetic elements are indicated and further described in Example 4G3).

Detailed Description of the Invention

A. Definitions

TABLE OF CORRESPONDENCE

Code	Group	Nucleotide(s)
A	A	adenine
C	C	cytosine
G	G	guanine
T	T	thymine (in DNA)
U	U	uracil (in RNA)
Y	C or T(U)	pyrimidine
R	A or G	purine
M	A or C	amino
K	G or T(U)	keto
S	G or C	strong interaction (3 hydrogen bonds)
W	A or T(U)	weak interaction (2 hydrogen bonds)
H	A or C or T(U)	not-G
B	G or T(U) or C	not-A

V	G or C or A	not-T or not-U
D	G or A or T(U)	not-C
N	G, A, C or T(U)	any

Amino Acid Residue: An amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide linkages. The amino acid residues described herein are preferably in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH_2 refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In keeping with standard polypeptide nomenclature (described in J. Biol. Chem., 243:3552-59 (1969) and adopted at 37 CFR §1.822(b)(2)), abbreviations for amino acid residues are shown in the following Table of Correspondence:

TABLE OF CORRESPONDENCE

SYMBOL		AMINO ACID
1-Letter	3-Letter	
Y	Tyr	tyrosine
G	Gly	glycine
F	Phe	phenylalanine
M	Met	methionine
A	Ala	alanine
S	Ser	serine
I	Ile	isoleucine
L	Leu	leucine

T	Thr	threonine
V	Val	valine
P	Pro	proline
K	Lys	lysine
H	His	histidine
Q	Gln	glutamine
E	Glu	glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	tryptophan
R	Arg	arginine
D	Asp	aspartic acid
N	Asn	asparagine
B	Asx	Asn and/or Asp
C	Cys	cysteine
X	Xaa	unknown/other

In addition the following have the meanings below:

BOC	tert-butyloxycarbonyl
DCCI	dicylcohexylcarbodiimide
DMF	dimethylformamide
OMe	methoxy
HOBt	1-hydroxybezotriazole

It should be noted that all amino acid residue sequences are represented herein by formulae whose left and right orientation is in the conventional direction of amino-terminus to carboxy-terminus. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino acid residues.

Polypeptide: A linear series of amino acid

residues connected to one another by peptide bonds between the alpha-amino group and carboxy group of contiguous amino acid residues.

Peptide: A linear series of no more than about 50 amino acid residues connected one to the other as in a polypeptide.

Protein: A linear series of greater than 50 amino acid residues connected one to the other as in a polypeptide.

Synthetic peptide: A chemically produced chain of amino acid residues linked together by peptide bonds that is free of naturally occurring proteins and fragments thereof.

Nucleotide: A monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is a nucleoside. When the nucleoside contains a phosphate group bonded to the 3' or 5' position of the pentose it is referred to as a nucleotide. A sequence of operatively linked nucleotides is typically referred to herein as a "base sequence" or "nucleotide sequence", and their grammatical equivalents, and is represented herein by a sequence whose left to right orientation is in the conventional direction of 5'-terminus to 3'-terminus.

Base Pair (bp): A partnership of adenine (A) with thymine (T), or of cytosine (C) with guanine (G) in a double stranded DNA molecule. In RNA, uracil (U) is substituted for thymine.

Nucleic Acid: A polymer of nucleotides, either single or double stranded.

Polynucleotide: A polymer of single or double stranded nucleotides. As used herein "polynucleotide" and its grammatical equivalents will include the full range of nucleic acids. A polynucleotide will typically refer to a nucleic acid molecule comprised of a linear strand of two or more deoxyribonucleotides and/or ribonucleotides. The exact size will depend on many factors, which in turn depends on the ultimate conditions of use, as is well known in the art. The polynucleotides of the present invention include primers, probes, RNA/DNA segments, oligonucleotides or "oligos" (relatively short polynucleotides), genes, vectors, plasmids, and the like.

Gene: A nucleic acid whose nucleotide sequence codes for an RNA or polypeptide. A gene can be either RNA or DNA.

Duplex DNA: A double-stranded nucleic acid molecule comprising two strands of substantially complementary polynucleotides held together by one or more hydrogen bonds between each of the complementary bases present in a base pair of the duplex. Because the nucleotides that form a base pair can be either a ribonucleotide base or a deoxyribonucleotide base, the phrase "duplex DNA" refers to either a DNA-DNA duplex comprising two DNA strands (ds DNA), or an RNA-DNA duplex comprising one DNA and one RNA strand.

Complementary Bases: Nucleotides that normally pair up when DNA or RNA adopts a double stranded configuration.

Complementary Nucleotide Sequence: A sequence of nucleotides in a single-stranded molecule of DNA or RNA that is sufficiently complementary to that on another single strand to specifically hybridize to it with consequent hydrogen bonding.

Recombinant DNA (rDNA) molecule: A DNA molecule produced by operatively linking two DNA segments. Thus, a recombinant DNA molecule is a hybrid DNA molecule comprising at least two nucleotide sequences not normally found together in nature. rDNA's not having a common biological origin, i.e., evolutionarily different, are said to be "heterologous".

Vector: A rDNA molecule capable of autonomous replication in a cell and to which a DNA segment, e.g., gene or polynucleotide, can be operatively linked so as to bring about replication of the attached segment. Vectors capable of directing the expression of genes encoding for one or more polypeptides are referred to herein as "expression vectors".

Conserved: A nucleotide sequence is conserved with respect to a preselected (reference) sequence if it non-randomly hybridizes to an exact complement of the preselected sequence.

Hybridization: The pairing of substantially complementary nucleotide sequences (strands of nucleic acid) to form a duplex or heteroduplex by the establishment of hydrogen bonds between complementary base pairs. It is a specific, i.e., non-random, interaction between two complementary polynucleotides that can be competitively inhibited.

Nucleotide Analog: A purine or pyrimidine

nucleotide that differs structurally from A, T, G, C, or U, but is sufficiently similar to substitute for the normal nucleotide in a nucleic acid molecule.

Upstream: In the direction opposite to the direction of DNA transcription, and therefore going from 5' to 3' on the noncoding strand, or 3' to 5' on the RNA transcript.

Downstream: Further along a DNA sequence in the direction of sequence transcription or read out, that is, traveling in a 3'- to 5'-direction along the noncoding strand of the DNA or 5'- to 3'-direction along the RNA transcript.

Stop Codon: Any of three codons that do not code for an amino acid, but instead cause termination of protein synthesis. They are UAG, UAA and UGA and are also referred to as a nonsense, termination, or translational stop codon.

Reading Frame: Particular sequence of contiguous nucleotide triplets (codons) employed in translation. The reading frame depends on the location of the translation initiation codon.

Homolog: Refers to a molecules that is structurally or functionally equivalent to a molecule of the present invention.

Fusion Protein: A polypeptide produced by recombinant DNA methods in which a first polypeptide domain is operatively linked to a second polypeptide domain by the peptide bond produced through expression of a single open reading frame to express a single "fused" polypeptide.

Chimeric Molecule: A bifunctional molecule formed

by connecting two separate molecules through chemical linkage, such as by crosslinking two isolated polypeptides or joining two heterologous fragments of DNA from different sources.

B. Translational Regulation by RB47 and RB60

1. RB47 and RB47 Binding Site

The present invention is based primarily on the discovery of that RB47, a mRNA binding protein, is a translational activator of the chloroplast *psb*A mRNA from *Chlamydomonas reinhardtii*, a green algae. The role of RB47 is now clearer, as evidence for its function comes from several independent sources. First, biochemical analysis has shown that this protein (along with other proteins within the complex) binds with high affinity and specificity to the *psb*A 5' UTR in a manner consistent with a role in translational regulation; high levels of binding are observed during high translation levels in the light, and low levels of binding are observed when translation is low in the dark. Second, the predicted amino acid sequence of RB47, now available since the cloning of RB47 as described herein, indicates the role this protein plays in translation. RB47 belongs to a family of proteins known as poly(A) binding proteins that bind RNA and have been shown to play a role in translation initiation (Bag and Wu, Eur. J. Biochem., 237:143-152 (1996); de Melo Neto et al., Nuc. Acids Res., 23:2198-2205 (1995); Proweller and Butler, J. Biol. Chem., 271:10859-10865 (1996); Sachs and Davis, Cell, 58:857-867 (1989); Tarun and Sachs, Embo J., 15:7168-7177 (1996)). Finally, genetic analysis has

predicted translational activators of chloroplast mRNAs (Girard-Bascou et al., Curr. Genet., 22:47-52 (1992); Kuchka et al., Embo J., 7:319-324 (1988); Rochaix et al., Ann. Rev. Cell Biol., 8:1-28 (1989); and Yohn et al., Mol. Cell. Biol., 16:3560-3566 (1996)). With the use and characterization of the hf149 mutant, a high fluorescence phenotype of *C. reinhardtii*, the absence of RB47 has been shown to correspond directly to the loss of translational initiation of the *psb*A mRNA, thus defining RB47 as a translational activator of the *psb*A mRNA. This is further supported by an additional nuclear mutation in *C. reinhardtii* (hf261) which is non-allelic to hf149, but shows the similar phenotype of a specific loss of D1 translation. The RB47 protein accumulates to less than 10% of the wild type level in this mutant.

While proteins which bind to the 5' UTR of chloroplast mRNAs seem likely candidates for translational activators, no direct link had been made to the body of genetic data prior to the characterization of the hf149 mutant as described in the present invention. Thus, the identification of RB47, the cloning thereof, and the role of RB47 in translation activation of *psb*A are novel and form the basis of the mechanisms of the expression cassettes of the present invention.

hf149 is not likely to be a mutation directly in the RB47 gene, as Southern and Northern analysis indicates that the RB47 gene is intact and produces normal amounts of RB47 mRNA in the hf149 strain. This leaves open the possibility that the loss of RB47

protein is the result of a loss of *psb*A translation, rather than the cause of it. Although this is a formal possibility, it is highly unlikely given the fact that the RB47 protein accumulates in other *psb*A translation initiation deficient mutants (e.g. F35, Yohn et al., Mol. Cell Biol., 16:3560-1566 (1996)), and that the *psb*A RB60 RNA binding protein still accumulates in the hf149 strain. Thus, the hf149 mutation provides strong evidence that the RB47 protein is directly involved in translational regulation of the chloroplast encoded *psb*A mRNA. Identification of the specific defect in the hf149 mutant should yield further insights into this process.

The dramatic reduction in the amount of *psb*A mRNA associated with ribosomes in the hf149 mutation suggests that RB47 is specifically required for ribosomes to initiate translation with the *psb*A mRNA. Although the identification of a message specific translational activator in the chloroplast has not previously been shown, other organellar systems may use similar mechanisms for controlling and coordinating gene expression, most notably the mitochondria of yeast. In particular, the *COX3* mRNA of *Saccharomyces cerevisiae* is translationally regulated by a complex of at least three proteins which have been shown genetically (Wiesenberger et al., Mol. Cell Biol., 15:3291-3300 (1995)) and biochemically (Brown et al., Mol. Cell Biol.,14:1045-1053 (1994)) to interact with each other and with the *COX3* mRNA. One of these proteins (PET122) also interacts with the mitochondrial ribosome (Haffter et al., Genetics, 127:319-326 (1991); Haffter et al.,

Genetics,125:495-503 (1990); McMullin et al., Mol. Cell Biol., 10:4590-4595 (1990)), suggesting a model for translational activation in which these proteins facilitate the initial interaction between the mRNA and the ribosome. A similar mechanism may be involved with RB47, the *psb*A mRNA and chloroplast ribosomes.

The identification of RB47 as a poly(A) binding protein (PABP) is somewhat unexpected given that translation in the chloroplast is generally considered prokaryotic like, and PABPs have not been identified as components of the prokaryotic translation apparatus. The chloroplast has 70S ribosomes (as in prokaryotes) and the mRNAs encoded by the chloroplast genome do not, in general, have poly(A) tails, and often contain prokaryotic consensus ribosome binding sequences (Gillham et al., Ann. Rev. Genetics, 28:71-93 (1994); Harris et al., Microbiol. Rev., 58:700-754 (1994)). The addition of A-rich sequences to the 3' end of endonucleolytic cleavage products of some chloroplast mRNAs has recently been described (Kudla et al., Embo J., 15:7137-7146 (1996); Lisitsky et al., Proc. Natl. Acad. Sci., USA, 93:13398-13403 (1996)), and this seems to play a role in degradation of the RNA, as in prokaryotes. The identification of a PABP in the chloroplast indicates that components of the cytoplasmic translation machinery may have been appropriated by the chloroplast for a similar function. These data also indicate that PABPs may function in translational regulation in the chloroplast in a manner not previously described for cytoplasmic mRNAs, although the role of RB47 in *psb*A translation seems to fit with the limited

information known about the function of PABPs in other systems. While no specific biochemical function has yet been identified for any member of the PABP family, these proteins have been defined as specific RNA binding proteins with a role in translational regulation. In yeast, PABP is essential for viability (Sachs et al., Cell, 45:827-835 (1986); Sachs et al., Mol. Cell Biol.,7:3268-3276 (1987)), and a temperature sensitive (ts) allele of PABP shows that depletion of PABP in yeast results in inhibition of translation initiation and poly(A) tail shortening (Sachs and Davis, Cell, 58:857-867 (1989)). Further, revertants of this ts mutation mapped to a ribosomal protein, suggesting that PABP interacts with the ribosome to mediate translation initiation (Sachs and Davis, id., (1989)). In addition, PABPs have been shown to physically interact with ribosomes (Proweller and Butler, supra, (1996)), and with eukaryotic initiation factors (eIF4G) (Tarun and Sachs, supra, (1996)). RB47 appears to fit these general roles predicted for PABPs, with the exception that RB47 shows specific binding to the 5' UTR of the *psb*A mRNA, and that RB47 is acting in the chloroplast, where translation is distinct from that in the cytoplasm both spatially and mechanistically. However, the fact that this nuclear encoded, eukaryotic protein has been exploited for use in the chloroplast may not be too surprising given the bi-directional exchange of genetic information between the chloroplast and nucleus (Morden et al., Biosystems, 28:75-90 (1992)).

Thus, in view of the binding specificity of RB47 to the RB47 binding site in *psb*A mRNA, the present

invention is unique in describing expression cassettes regulated at the translational level.

From the genetic, biochemical and molecular analysis of translational regulation in the chloroplast, a model for how the *psb*A mRNA binding proteins act in translation initiation and activation of *psb*A mRNA is presently formulated. In this model, nuclear encoded proteins, including the PABP homologue RB47, are transported into the chloroplast. Once in the plastid these proteins are activated to bind to RNA elements found within the 5' UTR of specific mRNAs (Mayfield et al., J. Cell Biol., 127:1537-1545 (1994)). This activation of binding is light responsive via the reducing potential generated by the light reactions of photosynthesis (Danon and Mayfield, Embo J., 13:2227-2235 (1994)). The interaction of the translational activator proteins and cis-acting RNA elements facilitates the initial interaction of the message with ribosomal subunits, resulting in increased translation of the D1 protein from the *psb*A mRNA.

Thus, in view of the enhancement of translation by the binding of a translation activator protein on a RNA element and in view of the cloning of both RB47 and RB60 translation activator proteins, the present invention contemplates the following aspects related to expression systems and uses thereof: isolated nucleic acids encoding recombinant proteins and variations thereof; the recombinant proteins themselves; use of the RNA binding site element in concert with the RB47 and RB60 nucleic acids and proteins thereof including endogenously expressed counterparts; expression

cassettes in which the genetic elements of this invention are operably linked; expression systems including cells *in vitro* and *in vivo*; methods of use thereof for expressing a heterologous molecule and for screening for agonists and antagonists of the interaction on which the present invention is based; and lastly, kits for use in expression of proteins and preparation of RNA transcripts.

The present invention therefore describes the use of an RB47 binding site nucleotide sequence and a coordinate RB47 binding site molecule for the purpose of enhancing translation of a desired heterologous coding sequence, thereby producing the desired expressed molecule for use thereafter.

Based on the translational activation mediated through the binding of an RB47 binding site sequence, typically a mRNA sequence, the present elements of RB47 binding site and an RB47 binding site polypeptide are therefore referred to as a translational activation system. The system is capable of further modulation or regulation by a polypeptide that regulates the binding of a RB47 binding site interaction with its activator protein as discussed in the next section. In a preferred embodiments, the translational activator protein of RB47 binding site is RB47 and the translation regulatory molecule that regulates the binding of an activator to an RB47 binding site is RB60, the latter of which is discussed below.

Thus, a molecule that binds RB47 and results in the translational activation of RB47 binding site thereby enhancing translation of a desired mRNA sequence is

referred to as a RB47 binding site polypeptide. Preferably, the polypeptide is RB47 that is present endogenously, i.e., naturally occurring, in a cell such that activation of a RB47 binding site occurs through an interaction of an endogenous protein with an exogenously provided RB47 binding site sequence as further described below. Functional RB47 protein is found in plant chloroplasts as reviewed above.

In other embodiments, RB47 is a recombinant protein produced through the expression of the coding sequence in a recipient cell as discussed in Section C below. Expression of a recombinant RB47 is now possible in view of the cloning of the cDNA encoding of RB47 as described in the present invention and in Examples 2 and 3. Exemplary recombinant RB47 proteins produced by the methods of this invention, more completely described in the Examples, include mature or processed RB47 that is approximately a 47 kilodalton (kDa) protein, precursor or unprocessed RB47 that is approximately a 69 kDa protein, and a histidine-modified RB47 protein that is also approximately a 47 kDa protein, the latter of which is useful for purification aspects as described in the Examples.

Thus, although the preferred RB47 proteins and nucleic acid compositions are derived from *Chlamydomonas reinhardtii* as discussed in the present invention, variations at both the amino acid and nucleotide sequence level may exist in similar functioning molecules isolated from different algae species as well as within differing plant geni. Such variations are not to be construed as limiting. For example, allelic

variation within a plant species can tolerate a several percent difference between isolates of a type of RB47, the differences of which comprise non-deleterious variant amino acid residues. Thus a protein of about 95% homology, and preferably at least 98% homology, to a disclosed RB47 protein is considered to be an allelic variant of the disclosed RB47 protein, and therefore is considered to be a RB47 protein of this invention.

Thus, the term "homolog" refers to any RB47-like protein or polypeptide having similar three-dimensional structure based on the amino acid residue sequence that can be encoded by differing specific nucleic acid sequences. In other words, the RB47 species of this invention are homologous molecules in view of the amino acid sequence similarity, the presence of a species specific sequence, the overall secondary and tertiary structure of the molecule, and the like physical parameters.

Thus as used herein, the phrases "RB47 protein" and "RB47 peptide or polypeptide" refers to a RB47 molecule having an amino acid residue sequence that comprises an amino acid residue sequence that corresponds, and preferably is identical, to a portion of a RB47 protein, either produced endogenously or exogenously to produce recombinant proteins, of this invention.

A recombinant RB47 protein need not necessarily be substantially pure, or even isolated, to be useful in certain embodiments, although recombinant production methods are a preferred means to produce a source for further purification to yield an isolated or substantially pure receptor composition. A recombinant

RB47 protein can be present in or on a mammalian cell line or in crude extracts of a mammalian cell line. In other embodiments, a recombinant RB47 protein is produced in or on plants or plant cell lines, for subsequent use therein to activate the translation of a desired coding sequence as described in Section C. Preferred expression vector systems for production of RB47 proteins of this invention in this context are described in Section C and in the Examples.

In the context of the present translational activation system of this invention, the presence of a RB47 binding site sequence is required. Thus, a RB47 binding site sequence is referred to as a translational activation binding domain, the activation of which leads to the enhanced or increased translation of a desired coding sequence. As previously discussed, endogenous activation of the RB47 mRNA binding sequence in the *psb*A gene in green algae chloroplast by RB47 results in the expression of D1 protein. As further discussed below, this translation activation can be further modulated or regulated by RB60. The nucleotide sequence in the 5' untranslated (5' UTR) end of the *psb*A gene containing the RB47 binding site is described in Example 3. The use of the RB47 binding site sequence thus is contemplated for use in preparing an expression cassette of this invention as further described in Section C and more completely in the Examples. Insofar as the binding of a RB47 binding site polypeptide to the nucleotide sequence of the RB47 binding site allows for translational activation of an encoding mRNA, variations, substitutions, additions, deletions and the

like permutations in the nucleic acid sequence of the RB47 nucleic acid sequence are contemplated for use in the present invention. In addition, any functional RB47 binding site nucleotide sequence is generally positioned upstream, i.e., 5', to the desired coding nucleotide sequence and in relation to the other inserted genetic elements including an upstream promoter, transcription initiation sites and downstream translation initiation sites of a coding region that can be a desired coding sequence or one of the genetic control elements of the invention, such as RB47 or RB60, as further described in Section C. and in the Examples.

2. RB60

Light-regulated translation of chloroplast mRNAs requires *trans*-acting factors that interact with the 5' untranslated region (UTR) of these mRNAs. The present invention describes a protein disulfide isomerase (PDI), also referred to as RB60, that is localized to the chloroplast and co-purifies with cPABP. The cDNA encoding the RB60 protein has now been cloned as described herein. As described more fully below, the RB60 protein has now been shown to modulate the binding of RB47, the cPABP, to the 5' UTR of the *psb*A mRNA by reversibly changing the redox status of cPABP using redox potential or ADP-dependent phosphorylation. This mechanism allows for a simple reversible switch regulating gene expression in the chloroplast. Moreover, in view of the modulatory properties of RB60 as discussed below, incorporation of RB60 into the compositions and methods of this invention are valuable

for regulating the expression of a desired gene product with the expression cassettes and systems of the present invention as described further herein and in the Examples.

The present inventors have determined the role of RB60 in regulating the binding of RB47 to *psb*A mRNA containing the RB47 binding site. The work has recently been published, Kim and Mayfield, Science, 278:1954-1957 (1997), references for the RB60 section herein are provided in the published paper. As previously discussed, synthesis of certain chloroplast photosynthetic proteins is activated 50-100 fold in response to light exposure without an increase in the corresponding mRNA levels, indicating that translation of chloroplast mRNAs is light-regulated. Genetic evidence has shown that nuclear-encoded *trans*-acting factors interact with the 5' untranslated region (UTR) of chloroplast mRNAs to activate translation of these mRNAs in a light-dependent fashion. A set of proteins (38, 47, 55 and 60 kDa) was identified to bind as a complex to the 5' UTR of the *psb*A mRNA, encoding the photosynthetic reaction center protein D1 from the green algae *Chlamydomonas reinhardtii*. Binding of this protein complex to the 5' UTR of the *psb*A mRNA correlates with light-enhanced translation of this mRNA under a variety of environmental condition, and in mutations deficient in *psb*A mRNA translation. RNA binding activity of the protein complex for the 5' UTR of the *psb*A mRNA can be regulated *in vitro* by at least two different mechanisms: ADP-dependent phosphorylation and changes in redox potential.

The present invention and the Kim and Mayfield, id., (1997) paper describe the cloning of the cDNA encoding the 60 kDa *psb*A mRNA binding protein (RB60) as further described in the Examples. The predicted amino acid sequence of the cloned cDNA is also described therein.

To verify that RB60 is localized to the chloroplasts, an immunoblot analysis of isolated pea chloroplasts was performed using the *C. reinhardtii* RB60 antiserum. To confirm that the isolated pea chloroplasts were free of cytoplasmic contamination, immunoblot analysis was performed with antiserum against the large subunit of ribulose bisphosphate carboxylase (RuBPCase, located in chloroplast) and antiserum against the cytoplasmic protein tubuli. RuBPCase antiserum recognized proteins from both whole leaf extracts (cytoplasm plus chloroplast) and from isolated chloroplasts. The tubulin antiserum recognized a protein in whole leaf extracts, but not in the chloroplast fraction), showing that the isolated chloroplasts were free of cytoplasmic proteins. The protein extracts from isolated pea chloroplasts were enriched using heparin-agarose chromatography: enrichment was required for immunoblot assays with the RB60 antiserum as RB60 is a minor component within the chloroplast. Immunoblot analysis was performed on proteins from purified pea chloroplasts, from *C. reinhardtii* cell extracts isolated by heparin-agarose chromatography, and on recombinant RB60. A specific signal immunochemically related to RB60 was clearly detected at approximately 63 kDa in the pea chloroplast

sample. A signal of equal intensity was observed for *C. reinhardtii* proteins and for the recombinant RB60.

Chloroplast PDI (cPDI) contains the two -CGHC- catalytic sites that are involved in the formation, reduction and isomerization of disulfide bonds associated with protein folding. The identification of these redox catalytic sites prompted the investigation of the role of RB60 in the redox-regulated binding of RB47 to the 5' UTR of the *psb*A mRNA. Both RB60 and RB47, containing only the four RNA recognition motif domains, were expressed as further described in the Examples in *E. coli* as a fusion protein with a histidine tag, purified on a Ni-NTA agarose affinity column and used for subsequent RNA binding gel mobility-shift assays. The effect of a reducing agent on RNA binding activity of recombinant RB47 (r-RB47) was assessed by the addition of DTT (dithiothreitol) in the presence of recombinant RB60 (r-RB60). r-RB47 was preincubated with 10 mM DTT, a 5-fold excess of r-RB60 alone, or both DTT plus r-RB60, prior to adding a ^{32}P-labeled 5'-UTR of the *psb*A mRNA, followed by a gel mobility-shift assay. The results showed that r-RB47 isolated from *E. coli* was in an active reduced form so that only a slight enhancement of RNA binding activity was obtained with addition of DTT and r-RB60.

To determine whether r-RB60 was able to re-activate r-RB47 that was in an inactive oxidized form, r-RB47 was incubated with the oxidant dithionitrobenzoic acid (DTNB) for 5 minutes and then dialyzed against 10^4 volume of buffer to remove the oxidant. Oxidation of r-RB47 by DTNB completely abolished the binding activity of the

protein. Addition of DTT to 1.0 mM partially restored the binding capacity of r-RB47, and the binding was increased three fold by the addition of up to 25 mM DTT. With increasing amounts of r-RB60, the binding activity of r-RB47 was increased compared to the samples without r-RB60 at every level of DTT tested. When DTT was not present in the incubation medium, r-RB60 alone could not restore the binding of the oxidized r-RB47 (0 mM DTT), indicating that r-RB60 requires reducing equivalents to convert the inactive oxidized form of r-RB47 to an active reduced form.

Protein disulfide isomerase is known to catalyze the formation of disulfide bonds by oxidation of the sulfhydryl groups of cysteine residues during protein folding. To examine whether r-RB60 was also capable of oxidative catalysis of the reduced form of r-RB47, GSSG, the oxidized form of the thiol tripeptide glutathionine, was added to the assay mixture. When GSSG alone was added to r-RB47 at up to 5 mM, there was a two fold reduction in binding activity of r-RB47 compared with untreated protein. Incubation of r-RB47 with both GSSG and r-RB60 reduced the binding activity of r-RB47 by 5-6 fold, indicating that r-RB60 can facilitate the conversion of the reduced form of r-RB47 to an inactive oxidized form under an oxidizing environment. Thus, RB60 modulates or in other words regulates the redox potential essential for RB47 binding activity. As such, RB60 is a regulatory protein useful in regulating the expression of a desired coding sequence in reducing and oxidizing environments as supported by the teachings described herein.

ADP-dependent phosphorylation of RB60 has previously been shown to reduce binding of the protein complex to the 5'-UTR of the *psb*A mRNA. To identify if recombinant RB60 can be phosphorylated, r-RB60 was incubated with heparin-purified proteins from *C. reinhardtii* in the presence of g-^{32}P-ATP. Phosphorylated r-RB60 was detected among a number of phosphorylated proteins in the heparin-purified fraction. Purification of the incubation mixtures on Ni-NTA resin resulted in the isolation of phosphorylated r-RB60. Phosphorylated r-RB60 was able to reduce the binding of r-RB47 to the 5' UTR of the *psb*A mRNA, whereas, phosphorylated *C. reinhardtii* proteins eluted from Ni-NTA resin had little impact on r-RB47 RNA binding.

It has previously been shown that thioredoxin can act as a transducer of redox potential to enhance the binding of a protein complex to the *psb*A mRNA. PDI fits well into this scheme as ferredoxin-thioredoxin reductase is capable of directly reducing PDI.

In a functional model of RB60 regulation, reducing equivalents, generated by photosynthesis, are donated to cPDI (RB60) through ferredoxin and ferredoxin-thioredoxin reductase and act to catalyze the reduction of chloroplast poly(A) binding protein (cPABP) (RB47). The reduced form of cPABP is then capable of binding to the 5' UTR of the *psb*A mRNA to activate translation initiation of this mRNA resulting in increased synthesis of the D1 protein. This mechanism provides a direct link in the chloroplast between the quantity of absorbed light and the rate of synthesis of the D1 protein, allowing the replacement of the

photo-damaged D1 protein. Protein disulfide isomerase has an additional advantage in this scheme in that it has greater oxidation potential than thioredoxin, thus allowing the off switch (oxidation) when reducing potential is low. ADP-dependent phosphorylation of RB60, that might be triggered by the increased pool of ADP during dark growth, can act to reduce the RNA binding activity of cPABP by enhancing the oxidative catalysis of cPDI over the reductive catalysis, resulting in decreased translation of the *psb*A mRNA. The data presented here show that a PDI such as RB60 acts as a regulator of RNA binding activity and hence gene expression, and not just as a catalyst for protein folding.

The present invention therefore describes the use of a protein disulfide isomerase, such as RB60, to function as a regulator of the binding of a RB47 binding polypeptide to the RB47 binding site nucleotide sequence for the activation of translation. Thus, in view of the foregoing disclosure, the use of a protein disulfide isomerase such as RB60 has many applicabilities in the context of the present invention, particularly ensuring translational control mechanisms for expression of a desired coding sequence and production of the encoded molecule in both oxidizing and reducing environments.

Based on the translational activation mediated through the binding of an RB47 binding site sequence, typically a mRNA sequence and the regulation by the additional element of translational regulator, the present elements of RB47 binding site, an RB47 binding site polypeptide and a RB60 or like molecule are

therefore referred to as a regulated translational activation system.

While the invention contemplates the use of any molecule that binds to RB47 binding site and any molecule that functions in accordance to the biological role of RB60 as described herein, in a preferred embodiments, the translational activator protein of RB47 binding site is RB47 and the translation regulatory molecule that regulates the binding of an activator to an RB47 binding site is RB60.

Preferably, a polypeptide that the regulates the binding of a separate polypeptide that binds to a RB47 binding site is present endogenously, i.e., naturally occurring, in a cell such that activation and regulation of translation mediated through a RB47 binding site occurs through an interaction of an endogenous protein with an exogenously provided RB47 binding site sequence as further described in Section C below and in the Examples. Functional RB60 protein is found endogenously found in plant chloroplasts as reviewed above.

In other embodiments, RB60 is a recombinant protein produced through the expression of the coding sequence in a recipient cell as discussed in Section C below. Expression of a recombinant RB60 is now possible in view of the cloning of the cDNA encoding of RB60 as described in the present invention and in Examples 2 and 3. An exemplary recombinant RB60 protein produced by the methods of this invention is more completely described in the Examples.

Thus, although the preferred RB60 protein and nucleic acid compositions are derived from *Chlamydomonas*

reinhardtii as discussed in the present invention, variations at both the amino acid and nucleotide sequence level may exist in similar functioning molecules isolated from different algae species as well as within differing plant geni. Such variations are not to be construed as limiting as previously discussed for RB47 compositions.

C. Recombinant DNA Molecules and Expression Systems that Utilize the RB47 Binding Site

The invention describes several nucleotide sequences of particular use in the methods of controlling gene expression using the RB47 binding site. These sequences include the actual RB47 binding site, the sequences which encode the RB47 protein that binds to the RB47 binding site, the RB60 protein which regulates the activity of RB47 protein, and various DNA segments, recombinant DNA (rDNA) molecules and vectors constructed for expression of these protein or for using these proteins to control expression of preselected structural genes.

DNA segments of this invention therefore can comprise sequences which encode whole structural genes, fragments of structural genes, and transcription units as described further herein.

A preferred DNA segment is a nucleotide sequence which defines an RB47 binding site as defined herein, which defines an RB47 protein, RB47 polypeptide or biologically active fragment thereof, or which defines an RB60 protein, RB60 polypeptide or biologically active fragment thereof.

The amino acid residue sequence of RB47 and of RB60 are described herein and in the Examples.

A preferred DNA segment codes for an amino acid residue sequence substantially the same as, and preferably consisting essentially of, an amino acid residue sequence or portions thereof corresponding to the RB47 or RB60 protein described herein. Representative and preferred DNA segments are further described in the Examples.

The amino acid residue sequence of a protein or polypeptide is directly related via the genetic code to the deoxyribonucleic acid (DNA) sequence of the structural gene that codes for the protein. Thus, a structural gene or DNA segment can be defined in terms of the amino acid residue sequence, i.e., protein or polypeptide, for which it codes.

An important and well known feature of the genetic code is its redundancy. That is, for most of the amino acids used to make proteins, more than one coding nucleotide triplet (codon) can code for or designate a particular amino acid residue. Therefore, a number of different nucleotide sequences may code for a particular amino acid residue sequence. Such nucleotide sequences are considered functionally equivalent since they can result in the production of the same amino acid residue sequence in all organisms. Occasionally, a methylated variant of a purine or pyrimidine may be incorporated into a given nucleotide sequence. However, such methylations do not affect the coding relationship in any way.

A nucleic acid is any polynucleotide or nucleic

acid fragment, whether it be a polyribonucleotide of polydeoxyribonucleotide, i.e., RNA or DNA, or analogs thereof. In preferred embodiments, a nucleic acid molecule is in the form of a segment of duplex DNA, i.e, a DNA segment, although for certain molecular biological methodologies, single-stranded DNA or RNA is preferred.

DNA segments are produced by a number of means including chemical synthesis methods and recombinant approaches, preferably by cloning or by polymerase chain reaction (PCR). DNA segments that encode portions of an RB47 or RB60 protein can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci et al, J. Am. Chem. Soc., 103:3185-3191, 1981, or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define the DNA segment, followed by hybridization and ligation of oligonucleotides to build the complete segment. Alternative methods include isolation of a preferred DNA segment by PCR with a pair of oligonucleotide primers.

Of course, through chemical synthesis, any desired modifications can be made simply by substituting the appropriate bases for those encoding the native amino acid residue sequence.

Furthermore, DNA segments consisting essentially of structural genes encoding an RB47 or RB60 protein can be subsequently modified, as by site-directed or random mutagenesis, to introduce any desired substitutions.

1. Cloning RB47 and RB60 Genes

An RB47 or RB60 gene of this invention can be

cloned by a variety of cloning methods using *Chlamydomonas reinhardtii* (*C. reinhardtii*) as a source of the genomic DNA or messenger RNA (mRNA) for cloning purposes. Cloning these genes can be conducted according to the general methods described in the Examples.

Preferred cloning strategies for isolating a nucleic acid molecule that encodes an RB47 or RB60 protein of this invention are described in the Examples.

Sources of libraries for cloning an RB47 or RB60 gene of this invention can include genomic DNA or messenger RNA (mRNA) in the form of a cDNA library from a tissue believed to express these proteins. The preferred tissue is plant chloroplast from *C. reinhardtii*.

A preferred cloning method involves the preparation a *C. reinhardtii* chloroplast cDNA library using standard methods, and preparing the RB47 or RB60-encoding nucleotide sequence using PCR with oligonucleotide primers based on the nucleotide sequences described herein for the RB47 or RB60 genes, respectively. Alternatively, the desired cDNA clones can be identified and isolated from a cDNA or genomic library by conventional hybridization methods using a hybridization probe based on the sequences described herein. Other methods are readily apparent to one skilled in the art.

2. Expression Vectors

In addition, the invention contemplates a recombinant DNA molecule (rDNA) containing a DNA segment of this invention encoding an RB47 or RB60 protein as described herein. A rDNA can be produced by operatively

(operably) linking a vector to a DNA segment of the present invention.

As used herein, the term "vector" refers to a DNA molecule capable of autonomous replication in a cell and to which another DNA segment can be operatively linked so as to bring about replication of the attached segment. A vector adapted for expression of a gene product and capable of directing the expression of a structural gene is referred to herein as an "expression vector". Thus, a recombinant DNA molecule is a hybrid DNA molecule comprising at least two nucleotide sequences not normally found together in nature.

The choice of vector to which a DNA segment of the present invention is operatively linked depends directly, as is well known in the art, on the functional properties desired, e.g., protein expression, and the host cell to be transformed, these being limitations inherent in the art of constructing recombinant DNA molecules. However, a vector contemplated by the present invention is at least capable of directing the replication, and preferably also expression, of a structural gene included in DNA segments to which it is operatively linked.

Both prokaryotic and eukaryotic expression vectors are familiar to one of ordinary skill in the art of vector construction and are described by Ausebel, et al., In Current Protocols in Molecular Biology, Wiley and Sons, New York (1993) and by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, (1989), which reference also describes all the general recombinant DNA methods

referred to herein.

In one embodiment, a vector contemplated by the present invention includes a procaryotic replicon, i.e., a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a procaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, those embodiments that include a procaryotic replicon also include a gene whose expression confers drug resistance to a bacterial host transformed therewith. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Those vectors that include a procaryotic replicon can also include a procaryotic promoter capable of directing the expression (transcription and translation) of a structural gene in a bacterial host cell, such as E. coli, transformed therewith. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA), pRSET available from Invitrogen (San Diego, CA) and pPL and pKK223 available from Pharmacia, Piscataway, N.J.

Expression vectors compatible with eucaryotic cells, preferably those compatible with vertebrate cells, can also be used to form the recombinant DNA

molecules of the present invention. Eucaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, #31255), pRc/CMV (Invitrogen, Inc.), the preferred vector pcDNA3 (Invitrogen) described in the Examples, and the like eucaryotic expression vectors.

An alternative expression system that can be used to express a protein of the invention is an insect system. In one such system, Autographa californica nuclear polyhidrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The polypeptide-encoding nucleotide sequence may be cloned into non-essential regions (in Spodoptera frugiperda for example the polyhedron gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedron promoter). Successful insertion of the polypeptide-encoding nucleotide sequence inactivates the polyhedron gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedron gene). These recombinant viruses are then used to infect cells in which the inserted gene is expressed. See Smith et al., J. Biol. Chem., 46:584 (1983); and Smith, U.S. Patent No. 4,215,051.

Mammalian cell systems that utilize recombinant viruses or viral elements to direct expression may be

engineered. For example, when using adenovirus expression vectors, the coding sequence of a polypeptide may be ligated to an adenovirus transcription/ translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted into the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the polypeptide in infected hosts (e.g., see Logan et al., Proc. Natl. Acad. Sci., USA, 81:3655-3659 (1984)). Alternatively, the vaccinia virus 7.5K promoter may be used. (e.g., see, Mackett et al., Proc. Natl. Acad. Sci., USA, 79:7415-7419 (1982); Mackett et al., J. Virol., 49:857-864 (1984); Panicali et al., Proc. Natl. Acad. Sci., USA, 79:4927-4931 (1982)). Of particular interest are vectors based on bovine papilloma virus which have the ability to replicate as extrachromosomal elements (Sarver et al., Mol. Cell. Biol., 1:486 (1981)). Shortly after entry of this DNA into mouse cells, the plasmid replicates to about 100 to 200 copies per cell. Transcription of the inserted cDNA does not require integration of the plasmid into the host's chromosome, thereby yielding a high level of expression. These vectors can be used for stable expression by including a selectable marker in the plasmid, such as the neo gene. Alternatively, the retroviral genome can be modified for use as a vector capable of introducing and directing the expression of the polypeptide-encoding nucleotide sequence in host cells (Cone et al., Proc. Natl. Acad. Sci., USA,

81:6349-6353 (1984)). High level expression may also be achieved using inducible promoters, including, but not limited to, the metallothionine IIA promoter and heat shock promoters.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with a cDNA controlled by appropriate expression control elements (e.g., promoter and enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. As mentioned above, the selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines.

For example, following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler et al., Cell, 11:223 (1977)), hypoxanthine-guanine phosphoribosyltransferase (Szybalska et al, Proc. Natl. Acad. Sci., USA, 48:2026 (1962)), and adenine phosphoribosyltransferase (Lowy et al., Cell, 22:817 (1980)) genes, which can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells respectively. Also, antimetabolite resistance-conferring genes can be used as the basis of selection; for example, the genes for dhfr, which confers resistance to methotrexate (Wigler

et al., Proc. Natl. Acad. Sci., USA, 77:3567 (1980); O'Hare et al., Proc. Natl. Acad. Sci., USA, 78:1527 (1981); gpt, which confers resistance to mycophenolic acid (Mulligan et al, Proc. Natl. Acad. Sci., USA, 78:2072, (1981)); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin et al, J. Mol. Biol., 150:1 (1981)); and hygro, which confers resistance to hygromycin (Santerre et al, Gene, 30:147 (1984)). Recently, additional selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman et al, Proc. Natl. Acad. Sci., USA, 85:804 (1988)); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue L., In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed., (1987)).

In another preferred embodiment, expression vectors compatible for use with plant cells are used to express structural genes in plants. Plants provide advantageous expression and delivery aspects in that a large supply of bulk protein with universal access is readily made from which the protein is either isolatable therefrom. Thus, transgenic plants containing expression vectors for encoding a recombinant protein of this invention is useful for preparing polypeptides of this invention.

Typical expression vectors useful for expression of genes in plants are well known in the art. Typical methods for introducing genes via expression vectors into plants include Agrobacterium tumefaciens-mediated

transformation, plant virus transfection, protoplast transformation, gene transfer into pollen, injection into reproductive organs, injection into immature embryos, and direct insertion, a process referred to as "biolystics". In the case of infection by plant viruses, a recombinant protein can be produced at high concentrations and isolated at low cost, with the genetic stocks being easily maintained for long periods of time without passaging through plants.

Preferred plants for such expression include any plant for which a compatible expression vector system exists, including dicots and monocots. Particularly preferred plants include alfalfa, tomato, petunia, soy bean, tobacco, corn, wheat, rice, spinach, asparagus, and the like. Exemplary plant expression vector systems for expression of a recombinant protein of this invention include those, such as binary vector system utilizing Agrobacterium tumefaciens, described in US Patent Number 5,202,422 and An et al., Plant Molecular Biology Manual, A3:1-19 (1988). Additional plant expression systems are described in US Patent Numbers 5,234,834, the disclosures of which several teachings are hereby incorporated by reference.

D. Methods

The present invention provides for a variety of methods using the disclosed recombinant proteins, nucleotide sequences, expression cassettes and expression systems. In particular, the invention provides methods for preparing (expressing) an RB47 or RB60 protein using the expression systems, methods for

controlling (regulating) the expression of a preselected coding sequence using the translation regulation properties of the RB47 system, methods for screening for useful agents which affect the ability of RB47 and/or RB60 to regulate translation in these systems, and the like methods.

1. Methods for Preparing a Recombinant Protein

An RB47 OR RB60 protein of this invention can be prepared by a variety of means, although expression using a rDNA expression vector is preferred. Exemplary production methods for a recombinant protein are described in the Examples.

In one embodiment, the invention contemplates methods for the preparation of a recombinant RB47 or RB60 proteins in their various forms using a nucleotide sequence-based expression system. The produced proteins are useful in the various embodiments described herein.

Although the description of expression is limited to specific examples, it is to be understood that the expression of proteins is generally characterized, and the expression of, for example, full length RB47 protein serves as an example of expression of any of a variety of forms of RB47 protein, including mature RB47, processed forms of RB47, biologically active fragments of RB47, fusion proteins containing RB47 domains, and the like. The descriptions herein apply to various forms of RB60 as well.

The method of preparing a recombinant RB47 or RB60 protein comprises providing an expression cassette as described herein that contains nucleotide sequences that

encode an RB47 or RB60 protein, or fragment thereof, together with nucleotide sequences that provide the requisite information for controlling gene expression and translation. The provided cassette is introduced into a suitable expression medium and maintained under conditions and for a time period sufficient for expression and translation of the protein product to occur. The times and conditions can vary, as is well known, depending upon the expression/translation medium (e.g., intracellular medium, *in vitro* expression medium, etc.). Nucleotide sequence information required for expression and translation are also well known in the art and need not be described in detail herein.

A typical expression system is described herein, in which the expression cassette is present on a recombinant plasmid that has been introduced into a microbial host. For example, the expression cassette is present in a PET expression plasmid introduced into *E. coli*, and the transformed bacterial is cultivated under growth conditions suitable for growth and expression of the expression cassette. Additional expression systems include other species of bacterial cells, yeast and eucaryotic cells, including mammalian cell expression systems, and *in vitro* expression systems, as are well known.

After expression, the expressed RB60 or RB47 protein is readily isolated from the expression medium (i.e., the host cell and cell contents) using standard biochemical separation methods to produce an isolated recombinant protein. Typical isolation methods can include disruption of the cell followed by protein

fractionations using mechanical, chemical, biological or immunological properties of the RB47 or RB60 protein. Preferred separation/isolation methods are described in the Examples.

Thus, the invention also provides a method for the production of recombinant proteins, either as intact RB47 or RB60 protein, as fusion proteins or as smaller polypeptide fragments of RB47 or RB60. The production method generally involves inducing cells to express a recombinant protein of this invention, recovering the expressed protein from the resulting cells, and purifying the expressed protein so recovered by biochemical fractionation methods, using a specific antibody of this invention, or other chemical procedures. Inducing expression of a recombinant protein can comprise inserting a rDNA vector encoding an RB47 or RB60 protein, or fragment thereof, of this invention, which rDNA is capable of expressing the structural gene encoding the RB47 or RB60 protein, into a suitable host cell, and expressing the vector's structural gene.

Thus, to facilitate expression of a recombinant protein or fusion protein of the present invention, DNA segments encoding either RB47 or RB60 as described herein, or portions thereof, are inserted into an expression vector. DNA segments are characterized as including a DNA sequence that encodes a recombinant protein of this invention, i.e., RB47 or RB60. That is, the DNA segments of the present invention are characterized by the presence of some or all of an RB47 or RB60 structural gene as described herein. Preferably

the gene is present as an uninterrupted linear series of codons where each codon codes for an amino acid residue found in the native protein, i.e., a gene free of introns.

2. Translational Regulation of Expression of a Coding Sequence

In a related embodiment, the invention contemplates methods for the controlled expression of a preselected coding sequence under the regulation of RB47 or a combination of RB47 and RB60 using the nucleotide sequences described herein that define an RB47 binding site and the recombinant proteins which bind this binding site and regulate translation of adjacent nucleotide sequences.

Thus, for example, the RB47 binding site can be engineered into an expression cassette as described herein to control the expression of a structural gene nucleotide sequence which encodes a preselected gene. The expression cassette contains the following genetic elements: (1) a promoter sequence that initiates transcription of a gene, (2) an RB47 binding site sequence adjacent to and located 3' relative to the promoter sequence, (3) a structural protein coding sequence under the expression control of the promoter, and a source of RB47 protein to regulate the expression of the cassette.

The promoter can be any of a variety of genetic elements as are well known for promotion of gene transcription. The promoter can be constitutive, inducible or repressible, thereby providing further

regulation. A preferred promoter is the lac z promoter inducible by IPTG, as is well known. Additional promoters include the T3 or T7 promoter.

The RB47 protein can be provided exogenously, as by addition of isolated protein to an expression medium containing the cassette, provided endogenously, as by introducing the cassette into a host cell which contains endogenous RB47 (e.g., a chloroplast cell that expresses RB47) or provided by introducing a gene which expresses the RB47 protein into the expression medium, either in combination with the cassette, or substantially contemporaneously with the cassette. The RB47 encoding gene can be added as a separate plasmid, or can be present as a second translation unit on the cassette which expresses the preselected gene.

Thus, in one embodiment, the invention comprises first forming an expression cassette by operably linking the above-identified components, and then introducing the expression cassette into a cell or other suitable expression medium.

Where the expression system can be further regulated by RB60, the RB60 protein can be added to the expression system exogenously from purified recombinant protein, provided as an endogenous protein when expression is carried out in a plant cell, or can be provided by expression from a second translation unit. The second transcription unit can be present on a separate nucleotide sequence, such as a separate plasmid capable of expressing RB60 that contains the RB60 nucleotide sequence, or present on the same expression cassette as a separate translation unit for RB60.

Expression cassettes can be introduced into an expression medium by any of a variety of means, and therefore the invention need not be so limited. For example, a variety of cell types can be used including bacterial, plant, yeast and higher eukaryote, all of which have different methods for transformation, including transduction, transfection, electroporation, transformation, biolistic bombardment, infection, and the like.

These systems provide particular advantages in the expression of preselected genes, including structural genes, insofar as these systems provide the ability to control timing and amounts of expression by specific and strong regulators of translation. The advantages will be apparent to one skilled in the art, but include synchronized expression in cell populations, combining expression with nutrient supplementation, regulated expression in therapeutic, manufacturing and diagnostic expression applications, and the like systems.

In one embodiment, the method for expression of a desired (preselected) coding sequence comprises first the method of preparing an expression cassette having the various components described herein, followed by introducing the cassette into an expression medium and maintaining the cassette under condition suitable for expression. To that end, the cassette can be prepared by any recombinant DNA method, which methods are well established in the art, including use of restriction enzymes to ligate nucleotide fragments, polymerase chain reactions (PCR) to isolate, mutate, modify and manipulate nucleotide fragments, and cloning sites for

insertion of preselected genes. An exemplary method involves operably linking the RB47 binding site sequence to a cloning site for insertion of a desired coding sequence, such that the cloning site is downstream of the binding site, and linking a second nucleotide sequence which encodes an RB47 polypeptide. The method can further involve linking a promoter 5' upstream to the RB47 binding site to form a transcription unit containing from 5' to 3' a promoter, a binding site and a cloning site for inserting the desired coding sequence. In a subsequent step for forming an expression cassette, the desired coding sequence is inserted into the cloning site.

Other permutations will be apparent to one skilled in the art.

3 Screening for Agonists and Antagonists of RB47-Mediated Translation

In another embodiment, the invention contemplates using an expression cassette containing an RB47 binding site to screen for agonists and antagonists which affect RB47 binding to the RB47 binding site, thereby identifying useful reagents for further control of an RB47-regulated (mediated) translation unit.

The method comprises providing an expression cassette according to the invention and having a indicator polypeptide as the desired structural gene into an expression system (i.e., medium), introducing RB47 and the candidate agent, and detecting the amount of indicator polypeptide expressed, and thereby the amount of effect the agent has on the expression system.

Controls are typically run in the presence and absence of the RB47 protein to demonstrate selectivity of the agent, which could be either an agonist or antagonist of RB47 activation of translation upon binding to the RB47 binding site.

Typical indicator polypeptides include enzymes which produce detectable substrates, light producing enzymes, such as luciferase, and the like. The RB47 can be added in the form of exogenous protein or by expression off of a nucleotide sequence, as described earlier.

In one embodiment, the expression system is a cell capable of supporting expression (transcription and translation) and the RB47 is provided in the cell either by adding protein to the cell or by providing a RB47-encoding nucleotide sequence to the cell.

In a further embodiment, the screening method is useful to identify agonists or antagonists of RB60 or RB60-mediated regulation of RB47-mediated translation, ie., reagents which effect RB60 rather than RB47 directly. This embodiment requires that the additional component RB60 be included in the screening method as described herein for expression using RB60.

Additional permutations are readily apparent to one skilled in the art.

E. Articles of Manufacture

The present invention also contemplates an article of manufacture comprising one or more of the components of the present invention. Typically, the article is present in the form of a package containing the

component or in combination with packaging material. The packaging material includes a label or instructions for use of the components of the package. The instructions indicate the contemplated use of the component as described herein for the methods or compositions of the invention.

For example, an article of manufacture can comprise one or both of the recombinant proteins of the invention, RB47 and RB60, in amounts useful in a method according to the invention. Alternatively, an article of manufacture can contain an expression cassette for expressing a desired coding sequence, which cassette comprises a construction as described herein that includes an RB47 binding site, and can optionally include a cloning site for insertion of a desired coding sequence, a promoter for controlling transcription of the expression cassette and inserted coding sequence, a coding sequence for the RB47 protein and/or the RB60 protein, and a preselected coding sequence. Alternatively, the article of manufacture may contain multiple nucleotide sequences, such as separate plasmid each encoding a different transcription, comprising one or more of the desired coding sequence under control of the RB47 binding site, the RB47 coding sequence and the RB60 coding sequence.

The article of manufacture may optionally include both an expression cassette and one or both of the recombinant proteins RB47 and RB60, or may contain a cell transformed by one or more of the expression cassettes of the present invention.

In a related embodiment, an expression cassette may

be used for expressing an RNA transcript containing an RB47 binding site, useful for subsequent regulation of translation of the transcript by RB47 protein. Such a construct can be used in the RNA expression field. Therefore, the invention contemplates an article of manufacture comprising packaging material, and in a separate container an expression cassette for expressing RNA that includes the RB47 binding site, wherein the packaging material includes a label that indicates the uses of the cassette in producing *in vitro* RNA transcripts. The production of RNA transcripts is well known. The article can further contain in separate containers components useful in combination with the cassette, including polymerases buffers, ribonucleotides and other reagents for *in vitro* transcription.

In these permutations, the components may optionally be present in the article of manufacture in separate containers.

Examples

The following examples relating to this invention are illustrative and should not, of course, be construed as specifically limiting the invention. Moreover, such variations of the invention, now known or later developed, which would be within the purview of one skilled in the art are to be considered to fall within the scope of the present invention hereinafter claimed.

1. Cloning of RB47

RB47 protein, 47 kilodalton (kDa) was purified by published procedures (Danon and Mayfield, Embo J.,

10:3993-4001 (1991)). The protein was then digested with proteinase Lys-C or trypsin, and the peptides separated HPLC and microsequenced (John Lesyk, Worchester Foundation for Experimental Biology, Worcester, MA and Arie Admon, The Protein Center, Department of Biology, Technion, Haifa, Israel). Two peptide sequences were obtained (QYGFVHFEDQAAADR (SEQ ID NO 1) and GFGFINFKDAESAA (SEQ ID NO 2)). Degenerate oligonucleotides were designed based on the reverse translation of these peptides. For the QYG... and GFG... peptides, the respective oligonucleotide sequences were 5'CAGTACGGYTTCGTBCAYTTCGAGGAYCAGGC3' (SEQ ID NO 3) and 5'GGAATTCGGYTTCGGYTTCATYAACTTCAAGGAYGCBGAG3' (SEQ ID NO 4), where the underline indicates an Eco RI restriction site and where Y=C or T; and B=G or T or C. A *C. reinhardtii* cDNA λ-gt10 phage library obtained from EMBL Laboratories, Heidelberg, Germany, was screened with these oligonucleotides using standard methods as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Chapter 8, 1989. One set of duplicate filter lifts was probed with each oligonucleotide, and plaques that hybridized to both were isolated. Several cDNA clones that hybridized to oligonucleotides from both peptides were identified. Four of these clones were 2.6 kilobases (kb) in length, the predicted full length of the RB47 mRNA. One of these cDNAs was subcloned into an *E. coli* plasmid for sequence determination using an automated sequencer.

The nucleotide and encoded amino acid sequence of

RB47 is also shown in Figures 1A-1D (SEQ ID NO 5). As described in Section 2 above, the predicted protein sequence from the cloned cDNA contained both the derived peptide sequences of RB47 and is highly homologous to poly(A) binding proteins (PABP) from a variety of eukaryotic organisms.

2. Cloning of RB60

To clone the cDNA encoding the 60 kDa *psb*A mRNA binding protein (RB60), the *psb*A-specific RNA binding proteins were purified from light-grown *C. reinhardtii* cells using heparin-agarose chromatography followed by *psbA* RNA affinity chromatography (RAC). RAC-purified proteins were separated by two-dimensional polyacrylamide gel electrophoresis. The region corresponding to RB60 was isolated from the PVDF membrane. RB60 protein was then digested with trypsin. Unambiguous amino acid sequences were obtained from two peptide tryptic fragments (WFVDGELASDYNGPR (SEQ ID NO 6) and (QLILWTTADDLKADAEIMTVFR (SEQ ID NO 7)) as described above for RB47. The calculated molecular weights of the two tryptic peptides used for further analysis precisely matched with the molecular weights determine by mass spectrometry. The DNA sequence corresponding to one peptide of 22 amino acid residues was amplified by PCR using degenerate oligonucleotides, the forward primer 5'CGCGGATCCGAYGCBGAGATYATGAC3' (SEQ ID NO 8) and the reverse primer 5'CGCGAATTCGTCATRATCTCVGCRTC3' (SEQ ID NO 9), where R can be A or G (the other IUPAC nucleotides have been previously defined above). The amplified sequence was then used to screen a λ-gt10 cDNA library

from *C. reinhardtii*. Three clones were identified with the largest being 2.2 kb. Selection and sequencing was performed as described for RB47 cDNA.

The resulting RB60 cDNA sequence is available via GenBank (Accession Number AF027727). The nucleotide and encoded amino acid sequence of RB60 is also shown in Figures 2A-2B (SEQ ID NO 10). The protein coding sequence of 488 amino acid residues corresponds to nucleotide positions 16-1614 of the 2413 base pair sequence. The predicted amino acid sequence of the cloned cDNA contained the complete amino acid sequences of the two tryptic peptides. The amino acid sequence of the encoded protein revealed that it has high sequence homology to both plant and mammalian protein disulfide isomerase (PDI), and contains the highly conserved thioredoxin-like domains with -CysGlyHisCys- (-CGHC-) (SEQ ID NO 11) catalytic sites in both the N-terminal and C-terminal regions and the -LysAspGluLeu- (-KDEL-) (SEQ ID NO 12) endoplasmic reticulum (ER) retention signal at the C-terminus found in all PDIs. PDI is a mutifunctional protein possessing enzymatic activities for the formation, reduction, and isomerization of disulfide bonds during protein folding, and is typically found in the ER. The first 30 amino acid residues of RB60 were found to lack sequence homology with the N-terminal signal sequence of PDI from plants or mammalian cells. However, this region has characteristics of chloroplast transit peptides of *C. reinhardtii*, which have similarities with both mitochondrial and higher plant chloroplast presequences. A transit peptide sequence should override the function

of the -KDEL- ER retention signal and target the protein to the chloroplast since the -KDEL- signal acts only to retain the transported protein in the ER.

3. Preparation of *psb*A Promoter Sequence and RB47 Binding Site Nucleotide Sequence

The chloroplast *psb*A gene from the green unicellular alga *C. reinhardii* was cloned and sequenced as described by Erickson et al., Embo J., 3:2753-2762 (1984), the disclosure of which is hereby incorporated by reference. The DNA sequence of the coding regions and the 5' and 3' untranslated (UTR) flanking sequences of the *C. reinhardii psb*A gene is shown in Figures 3A-3C. The *psb*A gene sequence is also available through GenBank as further discussed in Example 4. The nucleotide sequence is also listed as SEQ ID NO 13. The deduced amino acid sequence (also listed in SEQ ID NO 13) of the coding region is shown below each codon beginning with the first methionine in the open reading frame. Indicated in the 5' non-coding sequence are a putative Shine-Dalgarno sequence in the dotted box, two putative transcription initiation sites determined by S1 mapping (S1) and the Pribnow-10 sequence in the closed box. Inverted repeats of eight or more base pairs are marked with arrows and labeled A-D. A direct repeat of 31 base pairs with only two mismatches is marked with arrows labeled 31. Indicated in the 3' non-coding sequence is a large inverted repeat marked by a forward arrow and the SI cleavage site marking the 3' end of the mRNA. Both the 5' and 3' untranslated regions are used in preparing one of the expression cassettes of this

invention as further described below.

The 5' UTR as previously discussed contains both the *psb*A promoter and the RB47 binding site. The nucleotide sequence defining the *psb*A promoter contains the region of the *psb*A DNA involved in binding of RNA polymerase to initiate transcription. The -10 sequence component of the *psb*A promoter is indicated by the boxed nucleotide sequence upstream of the first S1 while the -35 sequence is located approximately 35 bases before the putative initiation site. As shown in Figure 3, the -10 sequence is boxed, above which is the nucleotide position (-100) from the first translated codon. The -35 sequence is determined accordingly. A *psb*A promoter for use in an expression cassette of this invention ends at the first indicated S1 site (nucleotide position -92 as counting from the first ATG) in Figure 3 and extends to the 5' end (nucleotide position -251 as shown in Figure 3). Thus, the promoter region is 160 bases in length. A more preferred promoter region extends at least 100 nucleotides to the 5' end from the S1 site. A most preferred region contains nucleotide sequence ending at the s1 site and extending 5' to include the -35 sequence, i.e., from -92 to -130 as counted from the first encoded amino acid residue (39 bases).

The *psb*A RB47 binding site region begins at the first S1 site as shown in Figure 3 and extends to the first adenine base of the first encoded methionine residue. Thus, a *psb*A RB47 binding site in the *psb*A gene corresponds to the nucleotide positions from -91 to -1 as shown in Figure 3.

The above-identified regions are used to prepare

expression constructs as described below. The promoter and RB47 binding site regions can be used separately; for example, the RB47 binding site sequence can be isolated and used in a eukaryotic or prokaryotic plasmid with a non-*psb*A promoter. Alternatively, the entire *psb*A 5' UTR having 251 nucleotides as shown in Figure 3 is used for the regulatory region in an expression cassette containing both the *psb*A promoter and RB47 binding site sequence as described below.

4. Preparation of Expression Vectors and Expression of Coding Sequences

A. Constructs Containing an *psb*A Promoter, an RB47 Binding Site Nucleotide Sequence, a Desired Heterologous Coding Sequence, an RB47-Encoding Sequence and an RB60-Encoding Sequence

Plasmid expression vector constructs, alternatively called plasmids, vectors, constructs and the like, are constructed containing various combinations of elements of the present invention as described in the following examples. Variations of the positioning and operably linking of the genetic elements described in the present invention and in the examples below are contemplated for use in practicing the methods of this invention. Methods for manipulating DNA elements into operable expression cassettes are well known in the art of molecular biology. Accordingly, variations of control elements, such as constitutive or inducible promoters, with respect to prokaryotic or eukaryotic expression systems as described in Section C.

are contemplated herein although not enumerated. Moreover, the expression the various elements is not limited to one transcript producing one mRNA; the invention contemplates protein expression from more than one transcript if desired.

As such, while the examples below recite one or two types of expression cassettes, the genetic elements of RB47 binding site, any desired coding sequence, in combination with RB47 and RB60 coding sequences along with a promoter are readily combined in a number of operably linked permeations depending on the requirements of the cell system selected for the expression. For example, for expression in a chloroplast, endogenous RB47 protein is present therefore an expression cassette having an RB47 binding site and a desired coding sequence is minimally required along with an operative promoter sequence. Overexpression of RB47 may be preferable to enhance the translation of the coding sequence; in that case, the chloroplast is further transformed with an expression cassette containing an RB47-encoding sequence. Although the examples herein and below utilize primarily the sequence encoding the precursor form of RB47, any of the RB47-encoding sequences described in the present invention, i.e., RB47 precursor, mature RB47 and histidine-modified RB47 are contemplated for use in any expression cassette and system as described herein. To regulate the activation of translation, an RB60-encoding element is provided to the expression system to provide the ability to regulate redox potential in the cell as taught in Section B. These examples herein and below

represent a few of the possible permutations of genetic elements for expression in the methods of this invention.

In one embodiment, a plasmid is constructed containing an RB47 binding site directly upstream of an inserted coding region for a heterologous protein of interest, and the RB47 and RB60 coding regions. Heterologous refers to the nature of the coding region being dissimilar and not from the same gene as the regulatory molecules in the plasmid, such as RB47 and RB60. Thus, all the genetic elements of the present invention are produced in one transcript from the IPTG-inducible *psb*A promoter. Alternative promoters are similarly acceptable.

The final construct described herein for use in a prokaryotic expression system makes a single mRNA from which all three proteins are translated. The starting plasmid is any *E. coli* based plasmid containing an origin of replication and selectable marker gene. For this example, the Bluescript plasmid, pBS, commercially available through Stratagene, Inc., La Jolla, CA, which contains a polylinker-cloning site and an ampicilin resistant marker is selected for the vector.

The wild-type or native *psb*A gene (Erickson et al., Embo J., 3:2753-2762 (1984), also shown in Figure 3, is cloned into pBS at the EcoRI and BamHI sites of the polylinker. The nucleotide sequence of the *psb*A gene is available on GenBank with the 5' UTR and 3' UTR respectively listed in Accession Numbers X01424 and X02350. The EcoRI site of *psb*A is 1.5 kb upstream of the *psb*A initiation codon and the BamHI site is 2 kb

downstream of the stop codon. This plasmid is referred to as pDl.

Using site-directed PCR mutagenesis, well known to one of ordinary skill in the art, an NdeI site is placed at the initiation codon of *psb*A in the pDl plasmid so that the ATG of the NdeI restriction site is the ATG initiation codon. This plasmid is referred to as pDl/Nde. An Nde site is then placed at the initiation codon of the gene encoding the heterologous protein of interest and an Xho I site is placed directly downstream (within 10 nucleotides) of the TAA stop codon of the heterologous protein coding sequence. Again using site-directed mutagenesis, an XhoI site is placed within 10 nucleotides of the initiation codon of RB47, the preparation of which is described in Example 2, and an NotI site is placed directly downstream of the stop codon of RB47. The heterologous coding region and the RB47 gene are then ligated into pDl/Nde so that the heterologous protein gene is directly adjacent to the RB47 binding site and the RB47 coding region is downstream of the heterologous coding region, using the Xho I site at the heterologous stop codon and the Not I site of the pDl polylinker.

These genetic manipulations result in a plasmid containing the 5' end of the *psb*A gene including the promoter region and with the RB47 binding site immediately upstream of a heterologous coding region, and the RB47 coding region immediately downstream of the heterologous coding region. The nucleotides between the stop codon of the heterologous coding region and the initiation codon of the RB47 coding region is preferably

less than 20 nucleotides and preferably does not contain any additional stop codons in any reading frame. This plasmid is referred to as pD1/RB47.

Using site-directed mutagenesis, a NotI site is placed immediately (within 10 nucleotides) upstream of the initiation codon of RB60, the preparation of which is described in Example 2, and an Xba I site is placed downstream of the RB60 stop codon. This DNA fragment is then ligated to the 3' end of the *psb*A gene using the Xba I site found in the 3' end of the *psb*A gene so that the *psb*A 3' end is downstream of the RB60 coding region. This fragment is then ligated into the pD1/RB47 plasmid using the NotI and BamHI sites so that the RB60 coding region directly follows the RB47 coding region. The resulting plasmid is designated pD1/RB47/RB60. Preferably there is less then 20 nucleotides between the RB47 and RB60 coding regions and preferably there are no stop codons in any reading frame in that region. The final plasmid thus contains the following genetic elements operably linked in the 5' to 3' direction: the 5' end of the *psb*A gene with a promoter capable of directing transcription in chloroplasts, an RB47 binding site, a desired heterologous coding region, the RB47 coding region, the RB60 coding region, and the 3' end of the *psb*A gene which contains a transcription termination and mRNA processing site, and an *E. coli* origin of replication and amplicillin resistance gene. A diagram of this plasmid with the restriction sites is shown in Figure 4.

Expression of pD1/RB47/RB60 in *E. coli* to produce recombinant RB47, RB60 and the recombinant heterologous

protein is performed as described in Example 4B. The heterologous protein is then purified as further described.

Expression cassettes in which the sequences encoding RB47 and RB60 are similarly operably linked to a heterologous coding sequence having the *psb*A RB47 binding site as described in Example 3 are prepared with a different promoter for use in eukaryotic, such as mammalian expression systems. In this aspect, the cassette is similarly prepared as described above with the exception that restriction cloning sites are dependent upon the available multiple cloning sites in the recipient vector. Thus, the RB47 binding site prepared in Example 3 is prepared for directed ligation into a selected expression vector downstream of the promoter in that vector. The RB47 and RB60 coding sequences are obtained from the pD1/RB47/RB60 plasmid by digestion with XhoI and XbaI and inserted into a similarly digested vector if the sites are present. Alternatively, site-directed mutagenesis is utilized to create appropriate linkers. A desired heterologous coding sequence is similarly ligated into the vector for expression.

B. Constructs Containing RB47 Nucleotide Sequence

1) Purified Recombinant RB47 Protein

In one approach to obtain purified recombinant RB47 protein, the full length RB47 cDNA prepared above was cloned into the *E. coli* expression vector pET3A (Studier et al., Methods Enzymol., 185:60-89 (1990)), also commercially available by Novagen,

Inc., Madison, WI and transformed into BL21 *E. coli* cells. The cells were grown to a density of 0.4 (OD_{600}), then induced with 0.5 mM IPTG. Cells were then allowed to grow for an additional 4 hours, at which point they were pelleted and frozen.

Confirmation of the identity of the cloned cDNA as encoding the authentic RB47 protein was accomplished by examining protein expressed from the cDNA by immunoblot analysis and by RNA binding activity assay. The recombinant RB47 protein produced when the RB47 cDNA was expressed was recognized by antisera raised against the *C. reinhardtii* RB47 protein. The *E. coli* expressed protein migrated at 80 kDa on SDS-PAGE, but the protein was actually 69 kDa, as determined by mass spectrometry of the *E. coli* expressed protein. This mass agrees with the mass predicted from the cDNA sequence. A 60 kDa product was also produced in *E. coli*, and recognized by the antisera against the *C. reinhardtii* protein, which is most likely a degradation or early termination product of the RB47 cDNA. The recombinant RB47 protein expressed from the RB47 cDNA is recognized by the antisera raised against the *C. reinhardtii* protein at levels similar to the recognition of the authentic *C. reinhardtii* RB47 protein, demonstrating that the cloned cDNA produces a protein product that is immunologically related to the naturally produced RB47 protein. In order to generate a recombinant equivalent of the endogenous native RB47, the location of the 47 kDa polypeptide was mapped on the full-length recombinant protein by comparing mass spectrometric data of tryptic digests of the *C. reinhardtii* 47 kDa protein and the

full-length recombinant protein. Thus, peptide mapping by mass spectrometry has shown that the endogenous RB47 protein corresponds primarily to the RNA binding domains contained within the N-terminal region of the predicted precursor protein, suggesting that a cleavage event is necessary to produce the mature 47 kDa protein. Thus, full-length recombinant RB47 is 69 kDa and contains a carboxy domain that is cleaved *in vivo* to generate the endogenous mature form of RB47 that is 47 kDa.

To determine if the heterologously expressed RB47 protein was capable of binding the *psb*A RNA, the *E. coli* expressed protein was purified by heparin agarose chromatography. The recombinant RB47 protein expressed in *E. coli* was purified using a protocol similar to that used previously for purification of RB47 from *C. reinhardtii*. Approximately 5 g of *E. coli* cells grown as described above were resuspended in low salt extraction buffer (10 mM Tris [pH 7.5], 10 mM NaCl, 10 mM $MgCl_2$, 5 mM ß-mercaptoethanol) and disrupted by sonication. The soluble cell extract was applied to a 5 mL Econo-Pac heparin cartridge (Bio-Rad) which was washed prior to elution of the RB47 protein (Danon and Mayfield, Embo J., 10:3993-4001 (1991)).

The *E. coli* expressed protein that bound to the heparin agarose matrix was eluted from the column at the same salt concentration as used to elute the authentic *C. reinhardtii* RB47 protein. This protein fraction was used in *in vitro* binding assays with the *psb*A 5' UTR. Both the 69 and 60 kDa *E. coli* expressed proteins crosslinked to the radiolabeled *psb*A 5' UTR at levels similar to crosslinking of the endogenous RB47 protein,

when the RNA/protein complex is subjected to UV irradiation.

Heparin agarose purified proteins, both from the *E. coli* expressed RB47 cDNA and from *C. reinhardtii* cells, were used in an RNA gel mobility shift assay to determine the relative affinity and specificity of these proteins for the 5' UTR of the *psb*A mRNA. The *E. coli* expressed proteins bound to the *psb*A 5' UTR *in vitro* with properties that are similar to those of the endogenous RB47 protein purified from *C. reinhardtii*. RNA binding to both the *E. coli* expressed and the endogenous RB47 protein was competed using either 200 fold excess of unlabeled *psb*A RNA or 200 fold excess of poly(A) RNA. RNA binding to either of these proteins was poorly competed using 200 fold excess of total RNA or 200 fold excess of the 5' UTR of the *psb*D or *psb*C RNAs. Different forms of the RB47 protein (47 kDa endogenous protein vs. the 69 kDa *E. coli* expressed protein) may account for the slight differences in mobility observed when comparing the binding profiles of purified *C. reinhardtii* protein to heterologously expressed RB47.

The mature form of RB47 is also produced in recombinant form by the insertion by PCR of an artificial stop codon in the RB47 cDNA at nucleotide positions 1403-1405 with a stop codon resulting in a mature RB47 recombinant protein having 402 amino acids as shown in Figure 1. An example of this is shown in Figure 5 for the production of a recombinant histidine-modified RB47 mature protein as described below. The complete RB47 cDNA is inserted into an expression

vector, such as pET3A as described above, for expression of the mature 47 kDa form of the RB47 protein. In the absence of the inserted stop codon, the transcript reads through to nucleotide position 2066-2068 at the TAA stop codon to produce the precursor RB47 having the above-described molecular weight characteristics and 623 amino acid residues.

Recombinant RB47 is also expressed and purified in plant cells. For this aspect, *C. reinhardtii* strains were grown in complete media (Tris-acetate-phosphate [TAP] (Harris, The *Chlamydonas* Sourcebook, San Diego, CA, Academic Press (1989)) to a density of $5x10^6$ cells/mL under constant light. Cells were harvested by centrifugation at 4°C for 5 minutes at 4,000 *g*. Cells were either used immediately or frozen in liquid N_2 for storage at -70°C.

Recombinant RB47 protein was also produced as a modified RB47 protein with a histidine tag at the amino-terminus according to well known expression methods using pET19-D vectors available from Novagen, Inc., Madison, WI. The nucleotide and amino acid sequence of a recombinant histidine-modified RB47 of the mature 47 kDa form is shown in Figure 5 with the nucleotide and amino acid sequence also listed in SEQ ID NO 14. Thus the nucleotide sequence of a histidine-modified RB47 is 1269 bases in length. The precursor form of the RB47 protein is similarly obtained in the expression system, both of which are modified by the presence of a histidine tag that allows for purification by metal affinity chromatography.

The recombinant histidine-modified RB47 purified

The recombinant histidine-modified RB47 purified through addition of a poly-histidine tag followed by Ni^{+2} column chromatography showed similar binding characteristics as that described for recombinant precursor RB47 described above.

C. Constructs Containing RB60 Nucleotide Sequence

In one approach to obtain purified recombinant RB60 protein, the full-length RB60 cDNA prepared above was cloned into the *E. coli* expression vector pET3A (Studier et al., Methods Enzymol., 185:60-89 (1990)), also commercially available by Novagen, Inc., Madison, WI and transformed into BL21 *E. coli* cells. The cells were grown to a density of 0.4 (OD_{600}), then induced with 0.5 mM IPTG. Cells were then allowed to grow for an additional 4 hours, at which point they were pelleted and frozen.

Recombinant histidine-modified RB60 was also expressed with a pET19-D vector as described above for RB47 that was similarly modified. Purification of the recombinant RB60 proteins was performed as described for RB47 thereby producing recombinant RB60 proteins for use in the present invention.

The RB60 coding sequence is also mutagenized for directional ligation into an selected vector for expression in alternative systems, such as mammalian expression systems.

D. Constructs Containing an RB47-Encoding Sequence and an RB60-Encoding Sequence

To prepare an expression cassette for encoding

both RB47 and RB60, one approach is to digest plasmid pD1/RB47/RB60 prepared above with XhoI and XbaI to isolate the fragment for both encoding sequences. The fragment is then inserted into a similarly digested expression vector if available or is further mutagenized to prepare appropriate restriction sites.

Alternatively, the nucleotide sequences of RB47 and RB60, as described in Example 2, are separately prepared for directional ligation into a selected vector.

An additional embodiment of the present invention is to prepare an expression cassette containing the RB47 binding site along with the coding sequences for RB47 and RB60, the plasmid pD1/RB47/RB60 prepared above is digested with NdeI and XhoI to prepare an expression cassette in which any desired coding sequence having similarly restriction sites is directionally ligated. Expression vectors containing both the RB47 and RB60 encoding sequences in which the RB47 binding site sequence is utilized with a different promoter are also prepared as described in Example 4A.

E. Constructs Containing an RB47 Binding Site Nucleotide Sequence, Insertion Sites for a Desired Heterologous Coding Sequence, and an RB47-Encoding Sequence

In another permutation, a plasmid or expression cassette is constructed containing an RB47 binding site directly upstream of an inserted coding region for a heterologous protein of interest, and the RB47 coding region. The final construct described herein for use in a prokaryotic expression system makes

a single mRNA from which both proteins are translated.

The plasmid referred to as pD1/RB47 is prepared as described above in Example 4A. A diagram of this plasmid with the restriction sites is shown in Figure 6.

Expression of pD1/RB47 in *E. coli* to produce recombinant RB47 and the recombinant heterologous protein is performed as described in above. The heterologous protein is then purified as further described.

To produce an expression cassette that allows for insertion of an alternative desired coding sequence, the plasmid pD1/RB47 is digested with NdeI and XhoI resulting in a vector having restriction endonuclease sites for insertion of a desired coding sequence operably linked to a RB47 binding site and RB47 coding sequence on one transcriptional unit.

F. Constructs Containing an RB47 Binding Site Nucleotide Sequence, Insertion Sites for a Desired Heterologous Coding Sequence, and an RB47-Encoding Sequence

In another permutation, a plasmid or expression cassette is constructed containing an RB47 binding site directly upstream of an inserted coding region for a heterologous protein of interest, and the RB60 coding region. The final construct described herein for use in a prokaryotic expression system makes a single mRNA from which both proteins are translated. In this embodiment, a separate construct encoding recombinant RB47 as described in Example 4B is co-transformed into the *E. coli* host cell for expression.

The plasmid referred to as pD1/RB60 is prepared as described above for pD1/RB47 in Example 4A with the exception that XhoI and XbaI sites are created on RB60 rather than RB47.

Expression of pD1/RB60 in *E. coli* to produce recombinant RB60 and the recombinant heterologous protein is performed as described in above with the combined expression of RB47 from a separate expression cassette. The heterologous protein is then purified as further described.

To produce an expression cassette that allows for insertion of an alternative desired coding sequence, the plasmid pD1/RB60 is digested with NdeI and XhoI resulting in a vector having restriction endonuclease sites for insertion of a desired coding sequence operably linked to a RB47 binding site and RB60 coding sequence on one transcriptional unit.

G. Constructs Containing RB47 Binding Site Nucleotide Sequence and Heterologous Coding Sequences

1) Expression of Recombinant Tetanus Toxin Single Chain Antibody

The examples herein describe constructs that are variations of those described above. The constructs described below contain an RB47 binding site sequence and a heterologous coding sequence. The activating protein RB47 was endogenously provided in the chloroplast and or plant cell. In other aspects however as taught by the methods of the present invention, the chloroplast is further transformed with an RB47-

expression construct as described above for overexpression of RB47 to enhance translation capacities.

A strain of the green algae *Chlamydomonas reinhardtii* was designed to allow expression of a single chain antibody gene in the chloroplast. The transgenically expressed antibody was produced from a chimeric gene containing the promoter and 5' untranslated region (UTR) of the chloroplast *psb*A gene prepared as described above, followed by the coding region of a single chain antibody (encoding a tetanus toxin binding antibody), and then the 3' UTR of the *psb*A gene also prepared as described above to provide for transcription termination and RNA processing signals. This construct is essentially pD1/Nde including a heterologous coding sequence having a 3' XbaI restriction site for ligation with the 3' *psb*A gene and is diagramed in Figure 7.

The *psb*A-single chain construct was first transformed into *C. reinhardtii* chloroplast and transformants were then screened for single chain gene integration. Transformation of chloroplast was performed via bolistic delivery as described in US Patents 5,545,818 and 5,553,878, the disclosures of which are hereby incorporated by reference. Transformation is accomplished by homologous recombination via the 5' and 3' UTR of the *psb*A mRNA.

As shown in Figure 8, two of the transformants that contained the single chain chimeric gene produced single chain antibodies at approximately 1% of total protein levels. The transgenic antibodies were of the correct

size and were completely soluble, as would be expected of a correctly folded protein. Few degradation products were detectable by this Western analysis, suggesting that the proteins were fairly stable within the chloroplast. To identify if the produced antibody retained the binding capacity for tetanus toxin, ELISA assays were performed using a mouse-produced Fab, from the original tetanus toxin antibody, as the control. The chloroplast single chain antibody bound tetanus toxin at levels similar to Fab, indicating that the single chain antibody produced in *C. reinhardtii* is a fully functional antibody. These results clearly demonstrate the ability of the chloroplast to synthesis and accumulate function antibody molecules resulting from the translational activation of an RB47 binding site in an expression cassette by endogenous RB47 protein in the chloroplast.

2) Expression of Bacterial Luciferase Enzyme Having Two Subunits

For the production of molecules that contain more than one subunit, such as dIgA and bacterial luciferase enzyme, several proteins must be produced in stoichiometric quantities within the chloroplast. Chloroplast have an advantage for this type of production over cytoplasmic protein synthesis in that translation of multiple proteins can originate from a single mRNA. For example, a dicistronic mRNA having 5' and 3' NdeI and XbaI restriction sites and containing both the A and B chains of the bacterial luciferase enzyme was inserted downstream of the *psbA* promoter and

5' UTR of the pD1/Nde construct prepared in Example 4A above. In this construct, the bacterial LuxAB coding region was ligated between the *psb*A 5' UTR and the *psb*A 3' end in an *E. coli* plasmid that was then transformed into *Chlamydomonas reinhardtii* cells as described above for expression in the chloroplast. A schematic of the construct is shown in Figure 9. Single transformant colonies were then isolated. A plate containing a single isolate was grown for 10 days on complete media and a drop of the luciferase substrate n-Decyl Aldehyde was placed on the plate and the luciferase visualized by video-photography in a dark chamber. Both proteins were synthesized from this single mRNA and luciferase activity accumulated within the chloroplast as shown in Figure 10. Some mRNA within plastids contained as many as 5 separate proteins encoded on a single mRNA.

3) Expression of Dimeric IgA

To generate dimeric IgA, the construct shown in Figure 11 is engineered so that the *psb*A promoter and 5' UTR are used to drive the synthesis of the light chain and heavy chains of an antibody, and the J chain normally associated with IgA molecules. The nucleic acid sequences for the dimeric IgA are inserted into the RB47 binding site construct prepared in Example 4A. The construct is then transformed into *C. reinhardtii* cells as previously described for expression of the recombinant dIgA.

Production of these three proteins within the plastid allows for the self assembly of a dimeric IgA (dIgA). Production of this complex is monitored in

several ways. First, Southern analysis of transgenic algae is used to identify strains containing the polycistronic chimeric dIgA gene. Strains positive for integration of the dIgA gene are screened by Northern analysis to ensure that the chimeric mRNA is accumulating. Western blot analysis using denaturing gels is used to monitor the accumulation of the individual light, heavy and J chain proteins, and native gels Western blot analysis will be used to monitor the accumulation of the assembled dIgA molecule.

By using a single polycistronic mRNA in the context of RB47 regulated translation, two of the potential pitfalls in the assembly of multimeric dIgA molecule are overcome. First, this construct ensures approximately stoichiometric synthesis of the subunits, as ribosomes reading through the first protein are likely to continue to read through the second and third proteins as well. Second, all of the subunits are synthesized in close physical proximity to each other, which increases the probability of the proteins self assembling into a multimeric molecule. Following the production of a strain producing dIgA molecules, the production of dIgA on an intermediate scale by growing algae in 300 liter fermentors is then performed. Larger production scales are then performed thereafter.

The foregoing specification, including the specific embodiments and examples, is intended to be illustrative of the present invention and is not to be taken as limiting. Numerous other variations and modifications can be effected without departing from the true spirit and scope of the invention.

What Is Claimed Is:

1. An expression cassette for expression of a desired molecule, which cassette comprises:

a) an RB47 binding site nucleotide sequence upstream of a restriction endonuclease site for insertion of a desired coding sequence to be expressed; and

b) a nucleotide sequence encoding a polypeptide which binds RB47 binding site.

2. The expression cassette of claim 1 further comprising a promoter sequence operably linked to and positioned upstream of the RB47 binding site nucleotide sequence.

3. The expression cassette of claim 2 wherein the promoter sequence is derived from a *psbA* gene.

4. The expression cassette of claim 3 wherein the coding sequence is heterologous to the *psbA* gene.

5. The expression cassette of claim 1 wherein the cassette comprises a plasmid or virus.

6. The expression cassette of claim 1 further comprising and operably linked thereto a nucleotide sequence encoding RB60.

7. The expression cassette of claim 1 wherein the RB47 binding polypeptide is selected from the group consisting of RB47, RB47 precursor and a histidine-modified RB47.

8. An expression cassette for expression of a desired molecule, which cassette comprises:

a) an RB47 binding site nucleotide sequence upstream of a restriction endonuclease site for insertion of a desired coding sequence to be expressed;

and

b) a nucleotide sequence encoding a polypeptide which regulates the binding of RB47 to the RB47 binding site.

9. The expression cassette of claim 8 wherein the regulatory polypeptide is RB60.

10. A recombinant RB47 protein.

11. A recombinant RB60 protein.

12. An isolated nucleotide sequence encoding RB47.

13. An isolated nucleotide sequence encoding a histidine-modified RB47.

14. An isolated nucleotide sequence encoding RB47 precursor.

15. The nucleotide sequence of claim 12 from nucleotide position 197 to 1402 in Figures 1A-1B and SEQ ID NO 5.

16. The nucleotide sequence of claim 13 from nucleotide position 1 to 1269 in Figures 5A-5B and SEQ ID NO 14.

17. The nucleotide sequence of claim 14 shown in from nucleotide position 197 to 2065 in Figures 1A-1C and SEQ ID NO 5.

18. An expression cassette comprising the nucleotide sequence of claim 12, 13 or 14.

19. An isolated nucleotide sequence encoding RB60.

20. The nucleotide sequence of claim 18 from nucleotide position 16 to 1614 in Figures 2A-2B and SEQ ID NO 10.

21. An expression cassette comprising the nucleotide sequence of claim 19.

22. An expression system comprising a cell

transformed with the expression cassette of claim 1.

23. The expression system of claim 22 wherein the cell is a plant cell.

24. The expression system of claim 23 wherein the plant cell endogenously expresses RB47.

25. The expression system of claim 23 wherein the plant cell endogenously expresses RB60.

26. The expression system of claim 23 wherein the plant cell endogenously expresses RB47 and RB60.

27. The expression system of claim 22 wherein the cell is a eukaryotic cell.

28. The expression system of claim 22 wherein the cell is a prokaryotic cell.

29. The expression system of claim 22 further comprising the expression cassette of claim 21.

30. An expression system comprising a cell transformed with the expression cassette of claim 8.

31. The expression system of claim 29 further comprising the expression cassette of claim 18.

32. A cell stably transformed with the expression cassette of claim 18.

33. A cell stably transformed with the expression cassette of claim 21.

34. A cell stably transformed with the expression cassette of claims 18 and 21.

35. The expression cassette of claim 1 further comprising an inserted desired coding sequence.

36. An expression system comprising a cell transformed with the expression cassette of claim 35, wherein the coding sequence is expressed forming the desired molecule upon activation of the RB47 binding

site with RB47.

37. The expression system of claim 36 wherein the cell is a plant cell endogenously expressing RB47.

38. The expression system of claim 36 wherein the cell is stably transformed with the expression cassette of claim 21.

39. An expression system comprising a cell transformed with an expression cassette comprising a promoter sequence, a RB47 binding site sequence, a desired coding sequence for a molecule, and a nucleotide sequence for encoding a polypeptide which binds RB47 binding site, wherein all sequences are operably linked.

40. A method of preparing a desired recombinant molecule wherein the method comprises cultivating the expression system of claim 36.

41. A method of preparing a desired recombinant molecule wherein the method comprises cultivating the expression system of claim 39.

42. A method for expressing a desired coding sequence comprising:

a) forming an expression cassette by operably linking:

1) a promoter sequence;

2) a RB47 binding site sequence;

3) a desired coding sequence; and

4) a nucleotide sequence encoding a polypeptide which binds RB47 binding site; and

b) introducing the expression cassette into a cell.

43. The method of claim 42 wherein the cell is a plant cell endogenously expressing RB47.

44. The method of claim 42 wherein the cell is a plant cell endogenously expressing RB60.

45. The method of claim 42 further comprising inducing expression with a promoter inducer molecule.

46. The method of claim 45 wherein the promoter inducer molecule is IPTG.

47. The method of claim 42 wherein the cell is transformed with the expression cassette of claim 21.

48. A method for expressing a desired coding sequence comprising:

a) forming an expression cassette by operably linking:

1) a promoter sequence;

2) a RB47 binding site sequence; and

3) a desired coding sequence;

and

b) introducing the expression cassette into a plant cell endogenously expressing RB47.

49. The method of claim 48 wherein the expression cassette further comprises a nucleotide sequence encoding RB60.

50. A method for the regulated production of a recombinant molecule from a desired coding sequence in a cell, wherein the cell contains the expression cassette of claim 34, wherein expression of the coding sequence is activated by RB47 binding to the RB47 binding site thereby producing the recombinant molecule.

51. A method of forming an expression cassette by operably linking:

a) a RB47 binding site sequence;

b) a cloning site for insertion of a desired

coding sequence downstream of the RB47 binding site sequence; and

c) a nucleotide sequence encoding a polypeptide which binds the RB47 binding site.

52. The method of claim 51 further comprising a promoter sequence operably linked upstream to the RB47 binding site sequence.

53. The method of claim 51 further comprising a desired coding sequence inserted into the insertion site.

54. A method of screening for agonists or antagonists of RB47 binding to RB47 binding site, the method comprising the steps:

a) providing a cell expression system containing:

1) a promoter sequence;

2) a RB47 binding site sequence;

3) a coding sequence for an indicator polypeptide; and

4) a polypeptide which binds to the RB47 binding site sequence;

b) introducing an antagonist or agonist into the cell; and

c) detecting the amount of indicator polypeptide expressed in the cell.

55. A method of screening for agonists or antagonists of RB60 in regulating RB47 binding to RB47 binding site, the method comprising the steps:

a) providing an expression system in a cell containing:

1) a promoter sequence;

2) a RB47 binding site sequence;

3) a coding sequence for an indicator polypeptide;

4) a polypeptide which binds to the RB47 binding site sequence; and

5) a RB60 polypeptide;

b) introducing an agonist or antagonist into the cell; and

c) detecting the amount of indicator polypeptide expressed in the cell.

56. An article of manufacture comprising a packaging material and contained therein in a separate container the expression cassette of claim 1, wherein the expression cassette is useful for expression of a desired coding sequence, and wherein the packaging material comprises a label which indicates that the expression cassette can be used for expressing a desired coding sequence when the RB47 binding site is activated by RB47.

57. The article of manufacture of claim 56 further comprising in a separate container the expression cassette of claim 18.

58. The article of manufacture of claim 56 further comprising in a separate container the expression cassette of claim 21.

59. An article of manufacture comprising a packaging material and contained therein in a separate container the expression system of claim 22, wherein the expression system is useful for expression of a desired coding sequence, and wherein the packaging material comprises a label which indicates that the expression

system can be used for expressing a desired coding sequence when the RB47 binding site is activated by RB47.

60. An article of manufacture comprising a packaging material and contained therein in a separate container the stably transformed cell of claim 32, wherein the cell is useful as an expression system, and wherein the packaging material comprises a label which indicates that the expression system can be used for expressing a desired coding sequence when the RB47 binding site is activated by RB47.

61. An article of manufacture comprising a packaging material and contained therein in a separate container the stably transformed cell of claim 33, wherein the cell is useful as an expression system, and wherein the packaging material comprises a label which indicates that the expression system can be used for expressing a desired coding sequence when the RB47 binding site is activated by RB47 and regulated by RB60.

62. An article of manufacture comprising a packaging material and contained therein in a separate container the stably transformed cell of claim 34, wherein the cell is useful as an expression system, and wherein the packaging material comprises a label which indicates that the expression system can be used for expressing a desired coding sequence when the RB47 binding site is activated by RB47 and regulated by RB60.

63. An article of manufacture comprising a packaging material and contained therein in a separate container the expression cassette of claim 2, wherein the expression cassette is useful for expression of a

RNA transcript, and wherein the packaging material comprises a label which indicates that the expression cassette can be used for producing *in vitro* a RNA transcript when the RB47 binding site is activated by RB47.

64. The article of manufacture of claim 63 wherein the promoter sequence is selected from the group consisting of T3 and T7 promoters.

65. The article of manufacture of claim 63 further comprising in separate containers a polymerase, a buffer and each of four ribonucleotides, reagents for *in vitro* RNA transcription.

```
  1 GAATTCGCGGCCGCTCCGTGGTTGGTCCTC ATG GTG TCT TTT TGA AGAGGACCTGAGCCTTTCACCCAAATATA 74
  1                                M   V   S   F   *                                5

 75 TCAAAAAACCCGGGCAACCGGCCAAAAAAATTGCAAAAGCCTCTCGTAGGCACAAAAGACCTATTCTAGCCATCAACTTT 154

155 GTATCCGACGCTGCCGTTTAGCTGCGCGTCTTGAAGTCAAGC ATG GCG ACT ACT GAG TCC TCG GCC CCG   223
  1                                           M   A   T   T   E   S   S   A   P     9

224 GCG GCC ACC ACC CAG CCG GCC AGC ACC CCG CTG GCG AAC TCG TCG CTG TAC GTC GGT GAC  283
 10 A   A   T   T   Q   P   A   S   T   P   L   A   N   S   S   L   Y   V   G   D    29

284 CTG GAG AAG GAT GTC ACC GAG GCC CAG CTG TTC GAG CTC TTC TCC TCG GTT GGC CCT GTG  343
 30 L   E   K   D   V   T   E   A   Q   L   F   E   L   F   S   S   V   G   P   V    49

344 GCC TCC ATT CGC GTG TGC CGC GAT GCC GTC ACG CGC CGC TCG CTG GGC TAC GCC TAC GTC  403
 50 A   S   I   R   V   C   R   D   A   V   T   R   R   S   L   G   Y   A   Y   V    69

404 AAC TAC AAC AGC GCT CTG GAC CCC CAG GCT GCT GAC CGC GCC ATG GAG ACC CTG AAC TAC  463
 70 N   Y   N   S   A   L   D   P   Q   A   A   D   R   A   M   E   T   L   N   Y    89

464 CAT GTC GTG AAC GGC AAG CCT ATG CGC ATC ATG TGG TCG CAC CGC GAC CCT TCG GCC CGC  523
 90 H   V   V   N   G   K   P   M   R   I   M   W   S   H   R   D   P   S   A   R   109

524 AAG TCG GGC GTC GGC AAC ATC TTC ATC AAG AAC CTG GAC AAG ACC ATC GAC GCC AAG GCC  583
110 K   S   G   V   G   N   I   F   I   K   N   L   D   K   T   I   D   A   K   A   129

584 CTG CAC GAC ACC TTC TCG GCC TTC GGC AAG ATT CTG TCC TGC AAG GTT GCC ACT GAC GCC  643
130 L   H   D   T   F   S   A   F   G   K   I   L   S   C   K   V   A   T   D   A   149

644 AAC GGC GTG TCG AAG GGC TAC GGC TTC GTG CAC TTC GAG GAC CAG GCC GCT GCC GAT CGC  703
150 N   G   V   S   K   G   Y   G   F   V   H   F   E   D   Q   A   A   A   D   R   169

704 GCC ATT CAG ACC GTC AAC CAG AAG AAG ATT GAG GGC AAG ATC GTG TAC GTG GCC CCC TTC  763
170 A   I   Q   T   V   N   Q   K   K   I   E   G   K   I   V   Y   V   A   P   F   189
```

FIG. 1A

```
764 CAG AAG CGC GCT GAC CGC CCC AGG GCA AGG ACG TTG TAC ACC AAC GTG TTC GTC AAG AAC  823
190 Q   K   R   A   D   R   P   R   A   R   T   L   Y   T   N   V   F   V   K   N    209

824 TTG CCG GCC GAC ATC GGC GAC GAC GAG CTG GGC AAG ATG GCC ACC GAG CAC GGC GAG ATC  883
210 L   P   A   D   I   G   D   D   E   L   G   K   M   A   T   E   H   G   E   I    229

884 ACC AGC GCG GTG GTC ATG AAG GAC GAC AAG GGC GGC AGC AAG GGC TTC GGC TTC ATC AAC  943
230 T   S   A   V   V   M   K   D   D   K   G   G   S   K   G   F   G   F   I   N    249

944 TTC AAG GAC GCC GAG TCG GCG GCC AAG TGC GTG GAG TAC CTG AAC GAG CGC GAG ATG AGC 1003
250 F   K   D   A   E   S   A   A   K   C   V   E   Y   L   N   E   R   E   M   S    269

1004 GGC AAG ACC CTG TAC GCC GGC CGC GCC CAG AAG AAG ACC GAG CGC GAG GCG ATG CTG CGC 1063
270  G   K   T   L   Y   A   G   R   A   Q   K   K   T   E   R   E   A   M   L   R    289

1064 CAG AAG GCC GAG GAG AGC AAG CAG GAG CGT TAC CTG AAG TAC CAG AGC ATG AAC CTG TAC 1123
290  Q   K   A   E   E   S   K   Q   E   R   Y   L   K   Y   Q   S   M   N   L   Y    309

1124 GTC AAG AAC CTG TCC GAC GAG GAG GTC GAC GAC GAC GCC CTG CGT GAG CTG TTC GCC AAC 1183
310  V   K   N   L   S   D   E   E   V   D   D   D   A   L   R   E   L   F   A   N    329

1184 TCT GGC ACC ATC ACC TCG TGC AAG GTC ATG AAG GAC GGC AGC GGC AAG TCC AAG GGC TTC 1243
330  S   G   T   I   T   S   C   K   V   M   K   D   G   S   G   K   S   K   G   F    349

1244 GGC TTC GTG TGC TTC ACC AGC CAC GAC GAG GCC ACC CGG CCG CCC GTG ACC GAG ATG AAC 1303
350  G   F   V   C   F   T   S   H   D   E   A   T   R   P   P   V   T   E   M   N    369

1304 GGC AAG ATG GTC AAG GGC AAG CCC CTG TAC GTG GCC CTG GCG CAG CGC AAG GAC GTG CGC 1363
370  G   K   M   V   K   G   K   P   L   Y   V   A   L   A   Q   R   K   D   V   R    389

1364 CGT GCC ACC CAG CTG GAG GCC AAC ATG CAG GCG CGC ATG GGC ATG GGC GCC ATG AGC CGC 1423
390  R   A   T   Q   L   E   A   N   M   Q   A   R   M   G   M   G   A   M   S   R    409
```

FIG. 1B

```
1424 CCG CCG AAC CCG ATG GCC GGC ATG AGC CCC TAC CCC GGC GCC ATG CCG TTC TTC GCT CCC 1483
 410 P   P   N   P   M   A   G   M   S   P   Y   P   G   A   M   P   F   F   A   P   429

1484 GGC CCC GGC GGC ATG GCT GCT GGC CCG CGC GCT CCG GGC ATG ATG TAC CCG CCC ATG ATG 1543
 430 G   P   G   G   M   A   A   G   P   R   A   P   G   M   M   Y   P   P   M   M   449

1544 CCG CCG CGC GGC ATG CCT GGC CCC GGC CGC GGC CCC CGC GGC CCC ATG ATG CCG CCC CAG 1603
 450 P   P   R   G   M   P   G   P   G   R   G   P   R   G   P   M   M   P   P   Q   469

1604 ATG ATG GGT GGC CCC ATG ATG GGC CCG CCC ATG GGC CCC GGG CGC GGC CGT GGC GGC CGC 1663
 470 M   M   G   G   P   M   M   G   P   P   M   G   P   G   R   G   R   G   G   R   489

1664 GGC CCC TCC GGC CGC GGC CAG GGC CGC GGC AAC AAC GCC CCT GCC CAG CAG CCC AAG CCC 1723
 490 G   P   S   G   R   G   Q   G   R   G   N   N   A   P   A   Q   Q   P   K   P   509

1724 GCC GCT GAG CCG GCC GCC GCG CCC GCC GCC GCC GCC CCC GCT GCC GCG GCG CCT GCC GCC 1783
 510 A   A   E   P   A   A   A   P   A   A   A   A   P   A   A   A   A   P   A   A   529

1784 GCG GCG GAG CCG GAG GCC CCC GCC GCC CAG CAG CCG CTG ACC GCC TCC GCG CTG GCC GCC 1843
 530 A   A   E   P   E   A   P   A   A   Q   Q   P   L   T   A   S   A   L   A   A   549

1844 GCC GCG CCG GAG CAG CAG AAG ATG ATG ATC GGC GAG CGC CTG TAC CCG CAG GTG GCG GAG 1903
 550 A   A   P   E   Q   Q   K   M   M   I   G   E   R   L   Y   P   Q   V   A   E   569

1904 CTG CAG CCC GAC CTG GCT GGC AAG ATC ACC GGC ATG CTG CTG GAG ATG GAC AAC GCC GAG 1963
 570 L   Q   P   D   L   A   G   K   I   T   G   M   L   L   E   M   D   N   A   E   589

1964 CTT CTG ATG CTT CTG GAG TCG CAC GAG GCG CTG GTG TCC AAG GTG GAC GAG GCC ATC GCT 2023
 590 L   L   M   L   L   E   S   H   E   A   L   V   S   K   V   D   E   A   I   A   609

2024 GTG CTC AAG CAG CAC AAC GTG ATT GCC GAG GAG AAC AAG GCT TAA AGCGCCTGCACGCTTGTGCG 2088
 610 V   L   K   Q   H   N   V   I   A   E   E   N   K   A   *                       624
```

FIG. 1C

```
2089 GGCTGGTGGCGCCGGCGCGCGCCGGCGCTGCTTGGGCCGCCGGCAGC ATG GGC GCG GCG GAC GCG GTG TGG  2159
   1                                                 M   G   A   A   D   A   V   W     8

2160 GAG CAG TGC TTG CTG CTT CTG GCC GCC GTG AAG CCG CGC CGA ACT GGG GCG GAC GGC AGG  2219
   9 E   Q   C   L   L   L   L   A   A   V   K   P   R   R   T   G   A   D   G   R    28

2220 CTG GCG TTG ACG CCG GCG CGC CAC AAC ACA AAG TTG GTG GCG TGA AAGTCTCTGGGCGTGCTCCG 2284
  29 L   A   L   T   P   A   R   H   N   T   K   L   V   A   *                       43

2285 GACGGTTGTAAGGTTTTAAGAACTGGCTTTTGGCCGGGTTGCCGCCCAAAGGCGGAACGGCGGTCTTTTCAGGCCAATCA 2364

2365 CATCCGGCTGGAAAAATTCTTACCAAAGCCAACCCCTGCACCCAAAAATTTCGGGTTCCGAAAGAACACTCCCCTTTTTT 2444

2445 CCGGCAACGCGTTCTTTCAAGGCCAATCACTTTCCGGGTTGGAAGAAA ATG TTA CCC GGA AAA GGC GGG AAG 2516
   1                                                 M   L   P   G   K   G   G   K     8

2517 CCC CCT GCA CCC GGA CAA GTT ATT CGG GGT TTC GCC GGG AAT GAG CAA GCG TTC GGG CTG  2576
   9 P   P   A   P   G   Q   V   I   R   G   F   A   G   N   E   Q   A   F   G   L    28

2577 TTG GCC GTA TCG CGA ACG CTG TCG GGG TGT CAG GCG CCA GAA GGA AGG ATG ACG TTT TGG  2636
  29 L   A   V   S   R   T   L   S   G   C   Q   A   P   E   G   R   M   T   F   W    48

2637 TGA AGGGGTGCAAACTGAGCACACGAGTTTTGGCAATAGACGTGGAGAAAGTCCAGTGCGGGGTGAGGCGGATAGCGGA 2715
  49 *                                                                                 49

2716 ATCAAGCGTGGCGGGTCCCTGGCGAGACGAGACGCTTCTGTTGTTTTGCTGAGCCCTTTG ATG GCA CAA TCG CAC 2790
   1                                                              M   A   Q   S   H     5

2791 TGT TTT GAG CAG GCG ACT GTA AAG TGC CCG ACG CTA AAA AAG CGG CCG CGA ATT CC       2846
   6 C   F   E   Q   A   T   V   K   C   P   T   L   K   K   R   P   R   I           23
```

FIG. 1D

MNRWNLLALTLGLLLVAAPFTKHQFAHASDEYEDDEEDDAPAAP

KDDDVDVTVVTVKNWDETVKKSKFALVEFYAPWCGHCKTLKPEYAKAATALKAAAPDA

LIAKVDATQEESLAQKFGVQGYPTLKWFVDGELASDYNGPRDADGIVGWVKKKTGPPA

VTVEDADKLKSLEADAEVVVVGYFKALEGEIYDTFKSYAAKTEDVVFVQTTSADVAKA

AGLDAVDTVSVVKNFAGEDRATAVLATDIDTDSLTAFVKSEKMPPTIEFNQKNSDKIF

NSGINKQLILWTTADDLKADAEIMTVFREASKKFKGQLVFVTVNNEGDGADPVTNFFG

LKGATSPVLLGFFMEKNKKFRMEGEFTADNVAKFAESVVDGTAQAVLKSEAIPEDPYE

DGVYKIVGKTVESVVLDETKDVLLEVYAPWCGHCKKLEPIYKKLAKRFKKVDSVIIAK

MDGTENEHPEIEVKGFPTILFYPAGSDRTPIVFEGGDRSLKSLTKFIKTNAKIPYELP
KKGSDGDEGTSDDKDKPASDKDEL

```
   1 gagtacgttt acgccatgaa ccgttggaac cttcttgccc ttaccctggg gctgctgctg
  61 gtggcagcgc ccttcaccaa gcaccagttt gctcatgctt ccgatgagta tgaggacgac
 121 gaggaggacg atgccccccgc cgcccctaag gacgacgacg tcgacgttac tgtggtgacc
 181 gtcaagaact gggatgagac cgtcaagaag tccaagttcg cgcttgtgga gttctacgct
 241 ccttggtgcg gccactgcaa gaccctcaag cctgagtacg ctaaggctgc caccgccctg
 301 aaggctgctg ctcccgatgc ccttatcgcc aaggtcgacg ccacccagga ggagtccctg
 361 gcccagaagt tcggcgtgca gggctacccc accctcaagt ggttcgttga tggcgagctg
 421 gcttctgact acaacggccc ccgcgacgct gatggcattg ttggctgggt gaagaagaag
 481 actggccccc ccgccgtgac cgttgaggac gccgacaagc tgaagtccct ggaggcggac
 541 gctgaggtcg ttgtcgtcgg ctacttcaag gccctggagg gcgagatcta cgacaccttc
 601 aagtcctacg ccgccaagac cgaggacgtg gtgttcgtgc agaccaccag cgccgacgtc
```

FIG. 2A

```
 661 gccaaggccg ccggcctgga cgccgtggac accgtgtccg tggtcaagaa cttcgccggt
 721 gaggaccgcg ccaccgccgt cctggccacg gacatcgaca ctgactccct gaccgcgttc
 781 gtcaagtcgg agaagatgcc ccccaccatt gagttcaacc agaagaactc tgacaagatc
 841 ttcaacagcg gcatcaacaa gcagctgatt ctgtggacca ccgccgacga cctgaaggcc
 901 gacgccgaga tcatgactgt gttccgcgag gccagcaaga agttcaaggg ccagctggtg
 961 ttcgtgaccg tcaacaacga gggcgacggc gccgaccccg tcaccaactt cttcggcctc
1021 aagggcgcca cctcgcctgt gctgctgggc ttcttcatgg agaagaacaa gaagttccgc
1081 atggagggcg agttcacggc tgacaacgtg gctaagttcg ccgagagcgt ggtggacggc
1141 accgcgcagg ccgtgctcaa gtcggaggcc atccccgagg acccctatga ggatggcgtc
1201 tacaagattg tgggcaagac cgtggagtct gtggttctgg acgagaccaa ggacgtgctg
1261 ctggagggtgt acgcccccctg gtgcggccac tgcaagaagc tggagcccat ctacaagaag
1321 ctggccaagc gctttaagaa ggtggattcc gtcatcatcg ccaagatgga tggcactgag
1381 aacgagcacc ccgagatcga ggtcaagggc ttccctacca tcctgttcta tcccgccggc
1441 agcgaccgca ccccatcgt gttcgagggc ggcgaccgct cgctcaagtc cctgaccaag
1501 ttcatcaaga ccaacgccaa gatcccgtac gagctgccca agaagggctc cgacggcgac
1561 gagggcacct cggacgacaa ggacaagccc gcgtccgaca aggacgagct gtaagcggct
1621 atctgaacta cccaggttt ggagcgtctg cttgcgcgct tgcgcttgca cactgtgcat
1681 ggatgggagt taaggaggag acggagcacg gaggctgcgc tcggttggtg gcttggagca
1741 ccggcagcgc gtgatccgtc ctggcagcag caacggcgga gcgggcgcat attggcgcga
1801 gctggcgagc ggctgttgct ggagagggata tgctgccggg cgggaggaag ggctaggggc
1861 agagatgaga gcgttacggg ctggcatgcg ggcgcccgtg cctctccctg cggtgcagtc
1921 cttgctagga gacgcacggt tttgccaaag agggacgctg tccacagccc tgcgactgga
1981 agttttttag gccctgcggt ggtagtggtg ttggtacggt tgtgtgcata agatgaacaa
2041 cgttctctc aagacgagac tactagtatg ctgacggtgt gtgtatgtgg tggatggatt
2101 gtgccccgac catgaagagt gctgtgttgc ctcggcgctt ctgtcgccct ggatgtgcgt
2161 ggttccgaac gctggagtca tctgttgagg agcgagggtg ttgtcgggtc cgcccggcac
2221 ggccgcgtga tgtccggatg gggattgcga gcgagggcaa ccgcagcgca gatagcgccg
2281 cagcggatcg agctagcgca ggatgatgag agccgggcct tcgcggcgtg ggatcaggga
2341 ggagccaagg cggagtgcat gcgaggaaaa cagtgtgcgg caaagaacgg gctgcaagaa
2401 cgccttgcgc aaa
```

FIG. 2B

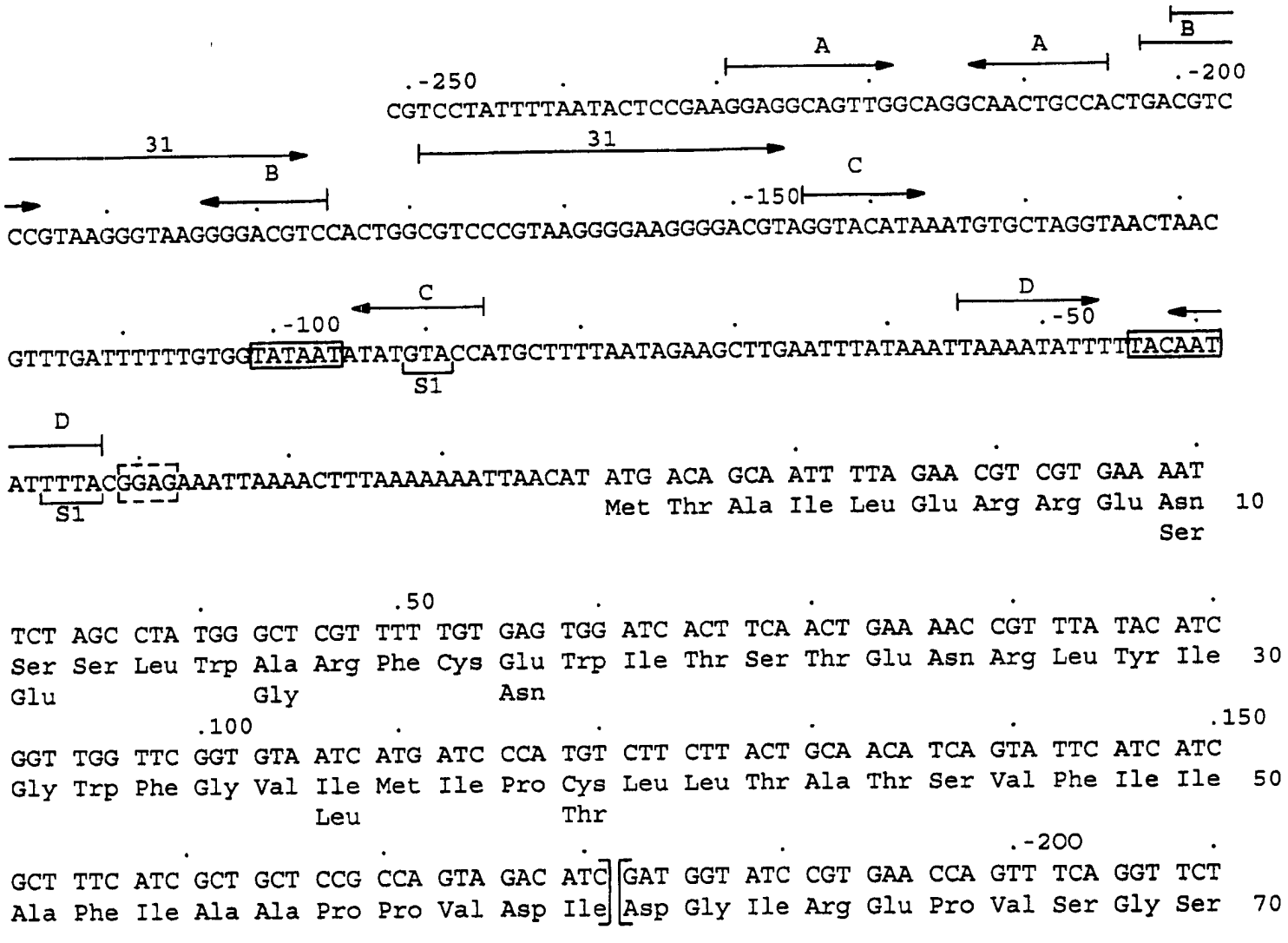


FIG. 3A

```
                                                    .250
CTT CTT TAC GGT AAC AAC ATC ATT ACA GGT GCT GTA ATC CCA ACT TCT AAC GCA ATC GGT
Leu Leu Tyr Gly Asn Asn Ile Ile Thr Gly Ala Val Ile Pro Thr Ser Asn Ala Ile Gly  90
                                Ser         Ile                 Ala

                                        .300
CTT CAC TTC TAC CCA ATT TGG GAA GCT GCT TCT CTA GAC GAG TGG TTA TAC AAC GGT GGT
Leu His Phe Tyr Pro Ile Trp Glu Ala Ala Ser Leu Asp Glu Trp Leu Try Asn Gly Gly 110
                                            Val

                        .350
CCT TAC CAA CTT ATC GTT TGT CAC TTC CTT CTA GGT GTA TAC TGC TAC ATG GGT CGT GAG
Pro Tyr Gln Leu Ile Val Cys His Phe Leu Leu Gly Val Tyr Cys Tyr Met Gly Arg Glu 130
      Glu               Leu                         Ala
            .400                                                    .450
TGG GAA TTA TCT TTC CGT TTA GGT ATG CGT CCA TGG ATC GCT GTA GCT TAC TCA GCT CCA
Trp Glu Leu Ser Phe Arg Leu Gly Met Arg Pro Trp Ile Ala Val Ala Tyr Ser Ala Pro 150

                                                        .500
GTA GCT GCA GCT TCA GCT GTA TTC TTA GTT TAC CCT ATC GGC CAA GGT TCA TTC TCT GAC
Val Ala Ala Ala Ser Ala Val Phe Leu Val Tyr Pro Ile Gly Gln Gly Ser Phe Ser Asp 170
               Thr                  Ile

                                    .550
GGT ATG CCT TTA GGT ATC TCT GGT ACT TTC AAC TTC ATG ATC GTA TTC CAA GCA GAA CAC
Gly Met Pro Leu Gly Ile Ser Gly Thr Phe Asn Phe Met Ile Val Phe Gln Ala Glu His 190

                                            .600
AAC ATC CTT ATG CAC CCA TTC CAC ATG TTA GGT GTT GCT GGT GTA TTC GGT GGT TCA TTA
Asn Ile Leu Met His Pro Phe His Met Leu Gly Val Ala Gly Val Phe Gly Gly Ser Leu 210

                        .850
TTC TCA GCT ATG CAC GGT TCT TTA GTT ACT TCA TCT TTA ATC CGT GAA ACA ACT GAA AAC
Phe Ser Ala Met His Gly Ser Leu Val Thr Ser Ser Leu Ile Arg Glu Thr Thr Glu Asn 230
```

FIG. 3B

```
                .700                                                .750
GAA TCA GCT AAC GAA GGT TAC CGT TTC GGT CAA GAA GAA GAA ACT TAC AAC ATT GTA GCT
Glu Ser Ala Asn Glu Gly Tyr Arg Phe Gly Gln Glu Glu Glu Thr Tyr Asn Ile Val Ala 250

                                                        .800
GCT CAT][GGT TAC TTT GGT CGT CTA ATC TTC CAA TAC GCT TCT TTC AAC AAC TCT CGT TCA
Ala His][Gly Tyr Phe Gly Arg Leu Ile Phe Gln Tyr Ala Ser Phe Asn Asn Ser Arg Ser 270

                                    .850
TTA CAC TTC TTC TTA GCT GCT TGG CCG GTA ATC GGT ATT TGG TTC ACT GCT TTA GGT TTA
Leu His Phe Phe Leu Ala Ala Trp Pro Val Ile Gly Ile Trp Phe Thr Ala Leu Gly Leu 290
                                    Val                                     Ile

                        .900
TCA ACT ATG GCA TTC AAC TTA AAC GGT TTC AAC TTC AAC CAA TCA GTA GTA GAC TCA CAA
Ser Thr Met Ala Phe Asn Leu Asn Gly Phe Asn Phe Asn Gln Ser Val Val Asp Ser Gln 310

                    .950
GGT CGT GTA CTA AAC ACT TGG GCA GAC ATC ATC AAC CGT GCT AAC TTA GGT ATG GAA GTA
Gly Arg Val Leu Asn Thr Trp Ala Asp Ile Ile Asn Arg Ala Asn Leu Gly Met Glu Val 330
            Ile                                                         .1050
            .1000
ATG CAC GAG CGT AAC GCT CAC AAC TTC CCT CTA GAC TTA GCT TCA ACT AAC TCT AGC TCA
Met His Glu Arg Asn Ala His Asn Phe Pro Leu Asp Leu Ala Ser Thr Asn Ser Ser Ser 350
                                            Ala Ile Glu Ala Pro
                                                    .1100
AAC AAC TAA TTT TTTTTTAAACTAAAATAAATCTGGTTAACCATACCTAGTTTATTTTAGTTTATACACACTTTT
                                                    S1
Asn Asn *Oc
Thr     Gly     *Oc
                .1150
CATATATATATACTTAATAGCTACCATAGGCAGTTGGCAGGACGTCCC
```

FIG. 3C

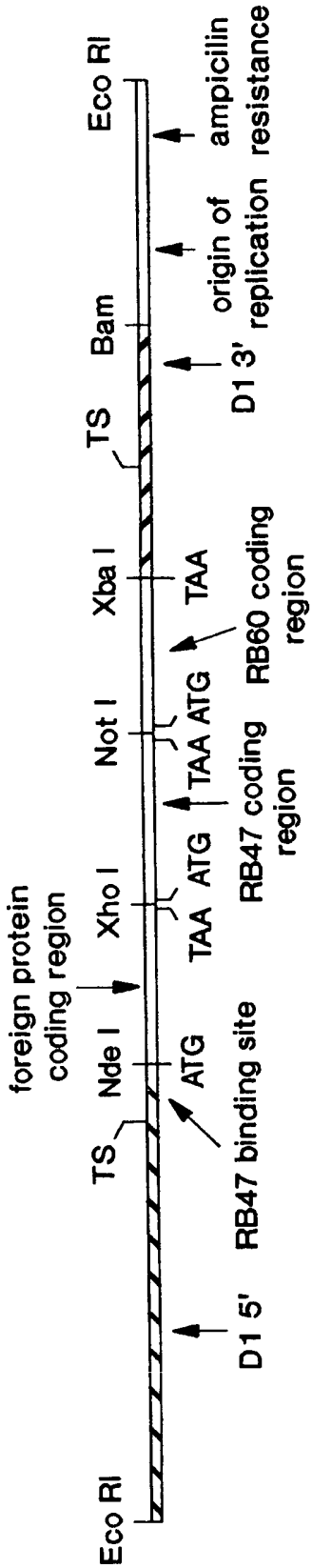

Eco RI
ampicilin resistance
origin of replication
D1 3'
Bam
TS
Xba I
TAA
RB60 coding region
Not I
TAA ATG
RB47 coding region
Xho I
TAA ATG
foreign protein coding region
Nde I
ATG
TS
RB47 binding site
D1 5'
Eco RI
TS = transcription start and transcription stop

FIG. 4

```
  1 ATG GGC CAT CAT CAT CAT CAT CAT CAT CAT CAT CAC AGC AGC GGC CAT ATC GAA GGT CGT  60
  1 M   G   H   H   H   H   H   H   H   H   H   H   S   S   G   H   I   E   G   R   20

 61 CAT ATG GCG ACT ACT GAG TCC TCG GCC CCG GCG GCC ACC ACC CAG CCG GCC AGC ACC CCG 120
 21 H   M   A   T   T   E   S   S   A   P   A   A   T   T   Q   P   A   S   T   P   40

121 CTG GCG AAC TCG TCG CTG TAC GTC GGT GAC CTG GAG AAG GAT GTC ACC GAG GCC CAG CTG 180
 41 L   A   N   S   S   L   Y   V   G   D   L   E   K   D   V   T   E   A   Q   L   60

181 TTC GAG CTC TTC TCC TCG GTT GGC CCT GTG GCC TCC ATT CGC GTG TGC CGC GAT GCC GTC 240
 61 F   E   L   F   S   S   V   G   P   V   A   S   I   R   V   C   R   D   A   V   80

241 ACG CGC CGC TCG CTG GGC TAC GCC TAC GTC AAC TAC AAC AGC GCT CTG GAC CCC CAG GCT 300
 81 T   R   R   S   L   G   Y   A   Y   V   N   Y   N   S   A   L   D   P   Q   A  100

301 GCT GAC CGC GCC ATG GAG ACC CTG AAC TAC CAT GTC GTG AAC GGC AAG CCT ATG CGC ATC 360
101 A   D   R   A   M   E   T   L   N   Y   H   V   V   N   G   K   P   M   R   I  120

361 ATG TGG TCG CAC CGC GAC CCT TCG GCC CGC AAG TCG GGC GTC GGC AAC ATC TTC ATC AAG 420
121 M   W   S   H   R   D   P   S   A   R   K   S   G   V   G   N   I   F   I   K  140

421 AAC CTG GAC AAG ACC ATC GAC GCC AAG GCC CTG CAC GAC ACC TTC TCG GCC TTC GGC AAG 480
141 N   L   D   K   T   I   D   A   K   A   L   H   D   T   F   S   A   F   G   K  160

481 ATT CTG TCC TGC AAG GTT GCC ACT GAC GCC AAC GGC GTG TCG AAG GGC TAC GGC TTC GTG 540
161 I   L   S   C   K   V   A   T   D   A   N   G   V   S   K   G   Y   G   F   V  180

541 CAC TTC GAG GAC CAG GCC GCT GCC GAT CGC GCC ATT CAG ACC GTC AAC CAG AAG AAG ATT 600
181 H   F   E   D   Q   A   A   A   D   R   A   I   Q   T   V   N   Q   K   K   I  200

601 GAG GGC AAG ATC GTG TAC GTG GCC CCC TTC CAG AAG CGC GCT GAC CGC CCC AGG GCA AGG 660
201 E   G   K   I   V   Y   V   A   P   F   Q   K   R   A   D   R   P   R   A   R  220
```

FIG. 5A

```
661 ACG TTG TAC ACC AAC GTG TTC GTC AAG AAC TTG CCG GCC GAC ATC GGC GAC GAC GAG CTG  720
221  T   L   Y   T   N   V   F   V   K   N   L   P   A   D   I   G   D   D   E   L  240

721 GGC AAG ATG GCC ACC GAG CAC GGC GAG ATC ACC AGC GCG GTG GTC ATG AAG GAC GAC AAG  780
241  G   K   M   A   T   E   H   G   E   I   T   S   A   V   V   M   K   D   D   K  260

781 GGC GGC AGC AAG GGC TTC GGC TTC ATC AAC TTC AAG GAC GCC GAG TCG GCG GCC AAG TGC  840
261  G   G   S   K   G   F   G   F   I   N   F   K   D   A   E   S   A   A   K   C  280

841 GTG GAG TAC CTG AAC GAG CGC GAG ATG AGC GGC AAG ACC CTG TAC GCC GGC CGC GCC CAG  900
281  V   E   Y   L   N   E   R   E   M   S   G   K   T   L   Y   A   G   R   A   Q  300

901 AAG AAG ACC GAG CGC GAG GCG ATG CTG CGC CAG AAG GCC GAG GAG AGC AAG CAG GAG CGT  960
301  K   K   T   E   R   E   A   M   L   R   Q   K   A   E   E   S   K   Q   E   R  320

961 TAC CTG AAG TAC CAG AGC ATG AAC CTG TAC GTC AAG AAC CTG TCC GAC GAG GAG GTC GAC 1020
321  Y   L   K   Y   Q   S   M   N   L   Y   V   K   N   L   S   D   E   E   V   D  340

1021 GAC GAC GCC CTG CGT GAG CTG TTC GCC AAC TCT GGC ACC ATC ACC TCG TGC AAG GTC ATG 1080
 341  D   D   A   L   R   E   L   F   A   N   S   G   T   I   T   S   C   K   V   M  360

1081 AAG GAC GGC AGC GGC AAG TCC AAG GGC TTC GGC TTC GTG TGC TTC ACC AGC CAC GAC GAG 1140
 361  K   D   G   S   G   K   S   K   G   F   G   F   V   C   F   T   S   H   D   E  380

1141 GCC ACC CGG CCG CCC GTG ACC GAG ATG AAC GGC AAG ATG GTC AAG GGC AAG CCC CTG TAC 1200
 381  A   T   R   P   P   V   T   E   M   N   G   K   M   V   K   G   K   P   L   Y  400

1201 GTG GCC CTG GCG CAG CGC AAG GAC GTG CGC CGT GCC ACC CAG CTG GAG GCC AAC ATG CAG 1260
 401  V   A   L   A   Q   R   K   D   V   R   R   A   T   Q   L   E   A   N   M   Q  420

1261 GCG CGC ATG TAA GGATCC                                                        1278
 421  A   R   M   *                                                                 424
```

FIG. 5B

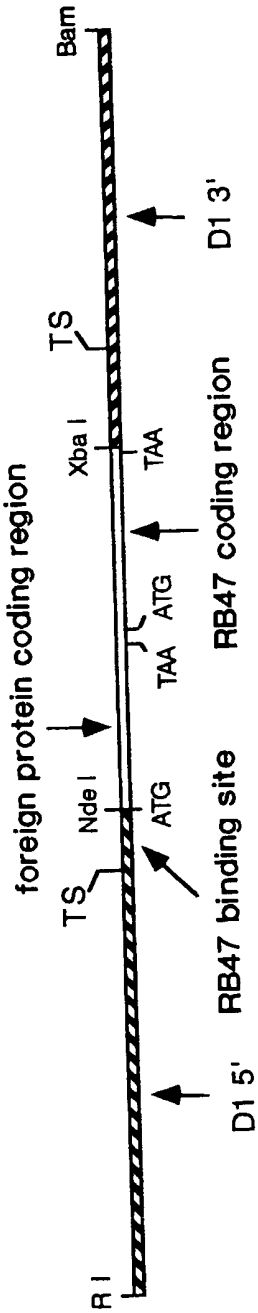

foreign protein coding region
Bam
TS
Xba I
TAA
D1 3'
Nde I
ATG
TAA
ATG
RB47 binding site
RB47 coding region
D1 5'
R I
TS = transcription start and transcription stop

FIG. 6

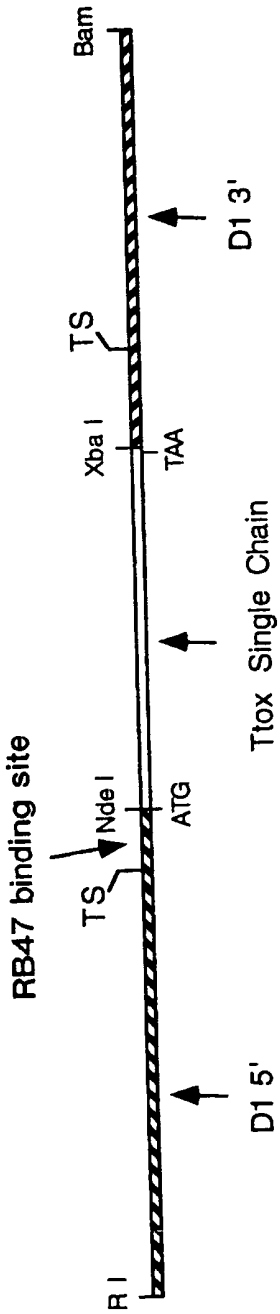

Bam
TS
D1 3'
Xba I
TAA
RB47 binding site
Nde I
ATG
Ttox Single Chain
or any coding region
TS
D1 5'
R I

FIG. 7

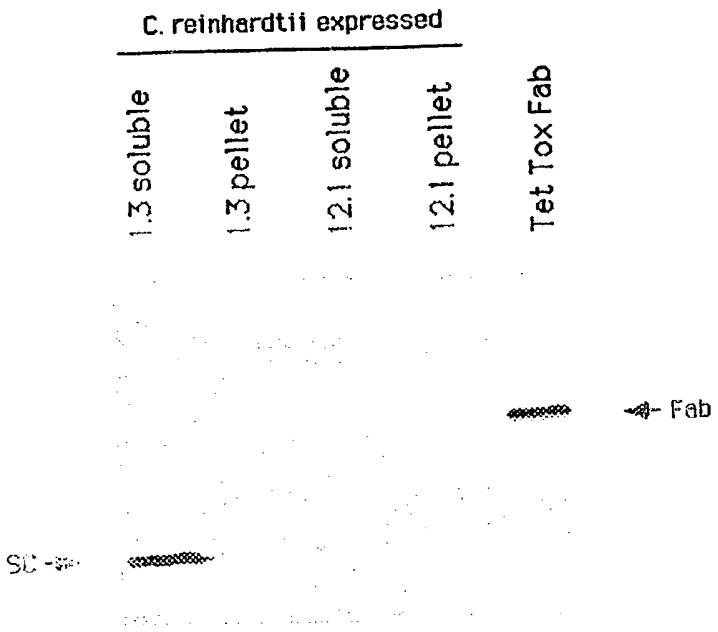
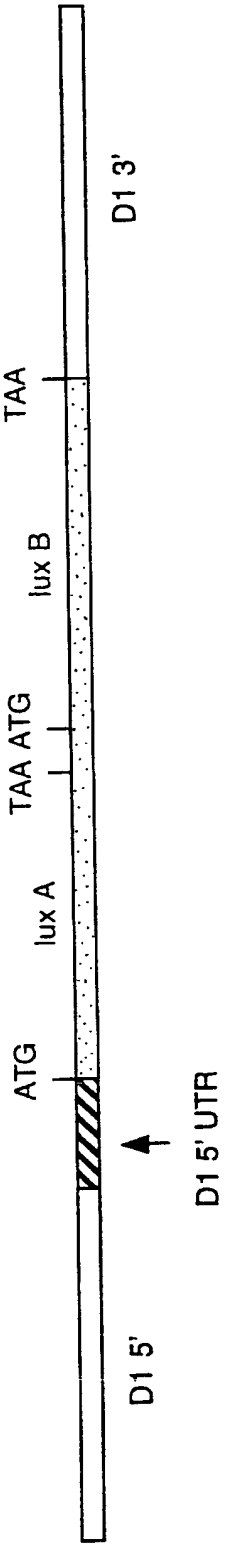

C. reinhardtii expressed
1.3 soluble
1.3 pellet
12.1 soluble
12.1 pellet
Tet Tox Fab
Fab
SC

FIG. 8



Bacterial luciferase A and B proteins expressed from a single mRNA containing the psbA 5' UTR with translational activator element.

FIG. 9

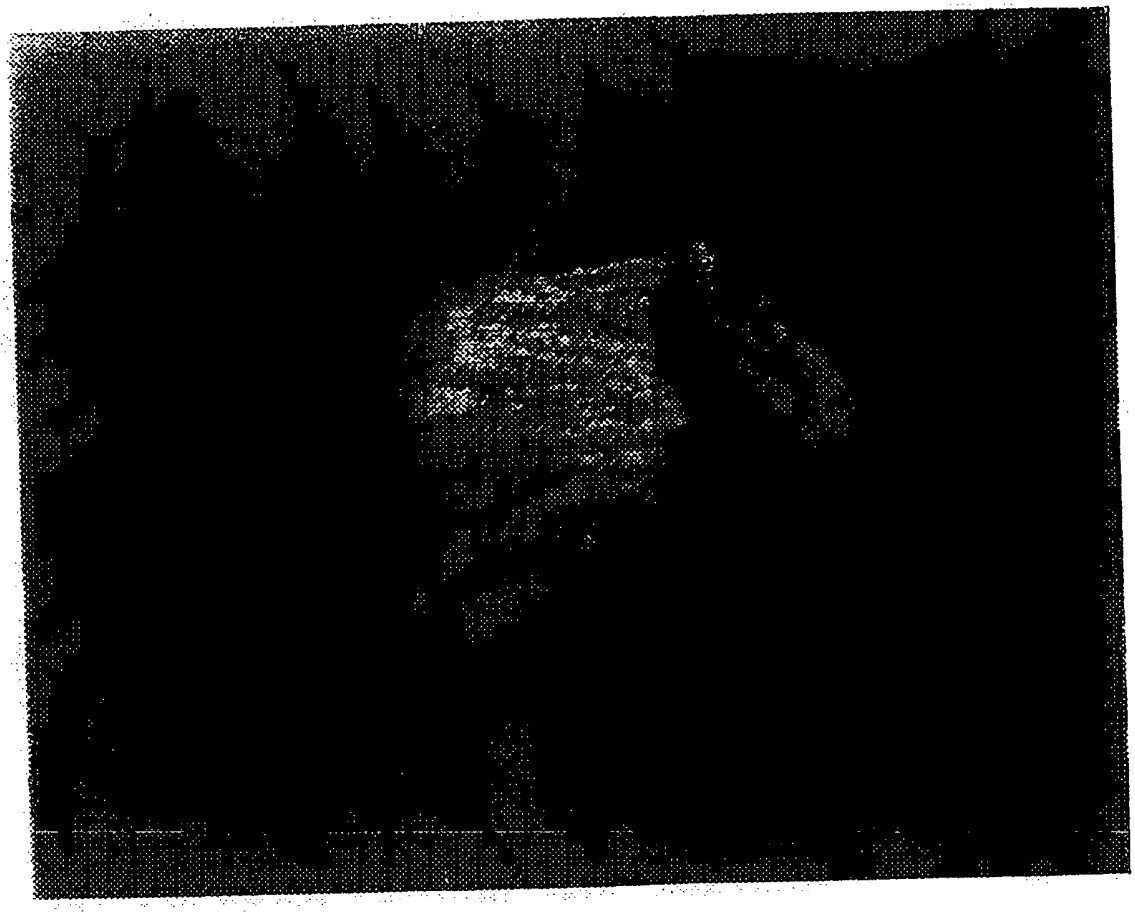


FIG. 10

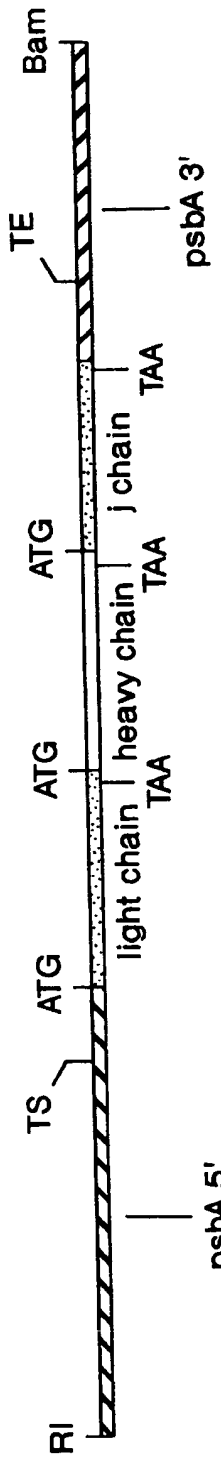

Bam
TE
psbA 3'
TAA
ATG
j chain
TAA
ATG
heavy chain
TAA
ATG
light chain
TS
psbA 5'
RI

FIG. 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00840

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12N 15/82, 5/10; C12P 21/02

US CL :435/69.1, 320.1, 325, 410; 530/350

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 320.1, 325, 410; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	YOHN. C.B. Altered mRNA Binding Activity and Decreased Translation Initiation in a Nuclear Mutant Lacking Translation of the Chloroplast psbA mRNA. Molecular and Cellular Biology. July 1996. Vol. 16. No. 7. pages 3560-3566, especially Figure 9.	10,11

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

23 APRIL 1998

Date of mailing of the international search report

23 JUN 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JAMES KETTER

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*